



Measurement of Online Education System in Pakistan during COVID-19: A Corpus Linguistic Analysis

Shafqat Rasool, Department of Education, Department Government College University Faisalabad, dr.shafqat.rasool@gcuf.edu.pk

Dr. Khuda Buksh Khan, Department of Education Department Govt. College University Faisalabad

Dr. Shumaila Shahzad, Department of Education Government College University Faisalabad

Urooj Fatima Alvi, Assistant Professor, English Department, University of Education Lahore

Abstract- The outbreak of the pandemic COVID-19 wobbled the entire educational system of the country analogous to the rest of the world. It leads to an abrupt eruption of online education. All the stakeholders are struggling to cope up with this sudden transformation. The higher education system of Pakistan is not an exception and is striving to survive against this global outburst. In the present research, the social media language used by various participants is explored through conducting multiple corpora and computational methodologies coupled with comparative analysis of survey results. The study focuses on the keyness, collocations, and visualization of the self-constructed corpus of electronic media. Multiple machine learning techniques (Sketch Engine, and Wordsmith version 8) have been used to retrieve the data and compile it into the corpus. The word 'online education has been observed regarding concordances, collocates, and n-grams. The common 4-gram phrase used throughout the corpus is 'the quality of education'. Moreover, survey results regarding students' online class experiences conducted from 8-11 June 2020 by a public sector university have been used as a reference for comparison. It shows that the survey results validate the text analysis outcomes. A total of 7198 students have responded to this online survey. Whereas a total of 37,796 words-based corpora based on online education, has been built up by the authors for this study. The hybrid methodological model of the current research encompasses comprehensive topographies of online learning and its influence on the community. It is valuable to contemplate the present research as seed work for exploring and quantifying the new dynamics of online education

Keywords: Pandemic COVID-19, Online Education, Sketch Engine, Wordsmith version 8, Survey Results

I. BACKGROUND OF THE RESEARCH ONLINE EDUCATION IN PAKISTAN

Previously, before the advent of pandemic COVID 19 crises, there were practically only ten universities providing online education /courses exclusively or moderately. These institutes include Virtual University of Pakistan (VU), Allama Iqbal Open University (AIOU), Preston University, Bahauddin Zakriya University (BZU), CASE-Centre of Advanced Studies in Engineering, University, University of Engineering and Technology (UET), Comsats University, University of Peshawar, Directorate of Distant Education City Campus, Gomal University in Dera Ismail Khan, Pakistan Institute of Modern Studies (PIMS) and Agha Khan University. Among these universities, The Virtual University of Pakistan provides an entirely online education based on modern information technologies. VU offers preeminent, cost-effective, and recognized courses on both national and international levels. Secondly, Allama Iqbal Open University, the oldest distance learning institute in Pakistan, provides the unsurpassed distance learning and technical education programs and courses to the masses (Ahsan, 2018).

Similarly, one of the oldest private higher education institute, Preston University, and its many campuses are involved in online education to the thousands of students with quality online knowledge. Other universities listed above are also taking part in radical online degree programs, advanced learning management systems, user grading, and communication systems. Moreover, these universities are also providing technical education programs for professional and part-time male and female students in different fields of engineering, information technology, and business administration through a virtual medium.

Although Allama Iqbal Open University introduced distance education first time in Asia and was second in the world, initially, it started the education system by hiring the tutors in the different institutes for its students nationwide. Later, adopting advanced information technology, the teaching method was converted from a traditional tutor system to the tutoring broadcast on television and radio network. The assessment procedure remained a regular paper-based examination and assignments. However, in 2002, the first IT-based University "Virtual University" was equipped with modern information technology

infrastructure, own television channel coupled with an online examination system (Abid et al., 2013, p. 111). Presently 174 universities, including both private and public, are operating under the auspices of Higher Education Commission of Pakistan (HEC). HEC is an independent and statutory body of primary funding, overseeing, regulating, and accrediting the higher education institution in Pakistan. It has the Motto "facilitating the institution of higher learning to serve as an engine of growth for Pakistan's socio-economic development (Higher Education Commission Pakistan,). Till 13th March 2020 all the universities, except one distance and one virtual university, were running customary face to face teaching with little or no online learning management system in place.

II. PANDEMIC COVID-19

Suddenly, the outbreak of the pandemic COVID-19 shook the entire education system of the country like all over the world. It leads to the eruption of online education suddenly. All the stakeholders are trying to cope up with this sudden transformation. Some of them had narrated this uncertain change and challenge through writing in the leading newspaper. In the table attached as Appendix A, salient newspaper articles have chosen to create a theoretical background of the study. As the pandemic crises are unforeseen and unprecedented, that is why the present research carries some delimitation, which is described after the problem statement.

Problem Statement

The outbreak of the pandemic COVID-19 has affected all walks of life. The higher education system of Pakistan is not an exception and is struggling against this global outburst by galvanizing the online education system to withstand and subsist. In this regard, the social media language used by various stakeholders is explored by conducting multiple corpus and computational methodologies. The study focuses on the keyness, collocations, and visualization of the corpus of electronic media for a better understanding of the readiness, challenges, and effects of the pandemic on the higher education sector of Pakistan.

Objectives

- i. To generate and identify the keywords and collocates for the online higher education system by scraping the official websites of four leading English newspapers of Pakistan.
- ii. To establish the critical outcomes of online education through detailed text analysis and multiple machine learning techniques (Sketch Engine, and Wordsmith version8).

Research Question

Primarily, the following questions address the phenomenon of online education overall.

1. To what extent the keyness and collocates are found and related to the extracted electronic media corpus to elaborate on the new change in the education system called online education.
2. Do the keywords and collocates show the actual state of mind of students, teachers, and other authoritative bodies.
3. Do the adopted techniques for text analysis provide the same outcome of the survey method?
4. Do the survey results align with the corpus methodology results?

Delimitations

1. For constructing the theoretical background, 12 newspaper articles published in the leading English newspaper articles have chosen for April, May, and June 2020 (Appendix A).
2. The corpus was compiled by scraping the official websites for four famous English newspapers.
3. The survey results for the comparisons of methodologies have used concerning one public sector university in which it was readily available and shared with the faculty forum. This survey was conducted to find the students' feedback.
4. The tools for the research, such as LancsBox and Sketch Engine, have been used after receiving the registration from copyrights.

III. THEORETICAL BACKGROUND

Pandemonium Online education has emerged with specific new experimentation in Pakistani higher education that should be examined and explored critically. It urges us to study the linguistic structure of online education narrative. To get a better understanding of the readiness, challenges, and effects of the pandemic on the higher education sector of Pakistan following are the major well-researched newspaper

articles that have been chosen to make the theoretical grounds of the study.

Firstly Khalid (2020) reported that Pandemic Covid-19 had brought opportunities and threats for the higher education sector. It has expedited the urge to go online and initiate diverse modes of delivery of higher education. But unequal access to the internet and electronic device to connect has hampered this effort to provide quality education to all students. Yet it has opened a new era of adopting new digital and online learning management systems. Collaboration amongst non-governmental organizations, the private sector, and government institutes can overcome this digital divide by providing internet access to all students and lending money to buy digital devices.

Similarly, Burhanuddin (2020) said that when all the universities were closed till 31st May 2020, starting from 13th March it was quite unusual and crises like situation for the management of higher education. Some universities were well equipped and prepared to go online and start from where the on-campus education was interrupted. Still, other universities do not have the infrastructure to launch online classes. Similarly, the geographical location of the university was playing a significant role in determining the speed of bandwidth and availability of the internet. Therefore, HEC has to devise flexible policy about all the universities, and a universal learning management system could not be implemented.

Mehmood (2020) expressed in his that Pakistan has many social, economic, political, and educational problems before the advent of COVID 19. This pandemic has worsened the situation and urged the HEC and Punjab HEC to come up with some solution to continue teaching and learning in the higher education institutes. The universities and HEC developed many new learning management systems. But already existing platforms i.e., zoom classroom, google classroom, and social media like Skype, WhatsApp, and YouTube, were used by different universities to provide multiple modes of learning to all students.

Batool (2020) reported that teaching in pandemic has changed from tradition face to face learning to new online education which can be termed as a Makeshift Teaching Model (MSTM) because it is not online learning in its real sense. Not only top leadership but also faculty members were not prepared to tackle an unprecedented situation and start a new mode of learning at once. On the other hand, the response of students was quite overwhelming, and the percentage of attendance has increased significantly. Students are the prime customers of the universities, and professors are trying to satisfy their customers. Still, it is counterproductive so far as coined by Jason Fertig "when professors strive for efficiency, learning decreases". Despite all these issues, online learning has a reasonable rate of return on investment as illustrated by Caroline Hoxby "though it provides a good growth of income, but not enough to make up for the cost of the education or even, in most cases, the cost to the individual student".

Hoodboy (2020) articulated his opinion that education plays a vital role in the socio- economic development of any country. In this period of the pandemic, training cannot be halted and wait to end this pandemic because there is no definite deadline for returning to normal. HEC has realized this grave situation and urged the universities to develop online and blended learning to involve all the students. Six vital components of online classes are the availability of the Learning Management System (LMS), faculty training, designing course for online learning, digitalization of the contents, willingness to start online learning, and, most importantly, involvement and readiness of the learners.

Alvi (2020) reported that many public sector universities have taken up the challenge of providing quality education to all the students by using asynchronous and synchronous online learning. Specially one public sector University i.e. the University of Education Lahore has formulated many committees to ensure quality of contents to be delivered to the students. Different applications like google meet, google classroom and zoom are used to deliver lectures and answer the queries of the students. A central repository on the YouTube is also established for asynchronous teaching.

Rehman (2020) reported that Pakistani graduates are relatively incompetent and having inferior logical and analytic skills even before the eruption of this pandemic. University faculty are teaching a few basic facts from textbooks and their prepared notes. Now when universities have started providing online education, the same ill-prepared lectures are being delivered to the students who can further worsen the already indisposed condition of higher education. But there is one silver lining that now anyone can compare and analyses the quality of contents and quality of teaching if lectures are available on the central repository.

Qazi (2020) reported that in Pakistan, there are 10.7 million boys and 8.6 million girls enrolled at the primary level. But not all students can continue their studies at the lower secondary level, and the number of enrolled children drops to 3.6 million boys and 2.8 million girls. Most of these students have limited access to the internet and smartphones. Now schools are sending online learning material, but the majority of the parents are illiterate and unable to teach or guide or even use a smartphone to help their children to study online. The government has taken many steps to overcome this situation, i.e., National Knowledge Bank, Technical Support Committee, to cater to hybrid mode and launching of the-school.

Bari (2020) Education has never been among the top priorities of the Government of Pakistan. This has been reflected in the confusing and changing stance of the Government about online education. So far there have been six different policy notifications have been issued. It all started with declaring early summer vacations till 15th July 2020 and then asking the universities to start online classes. Universities have blindly followed the revolutions of the HEC to get high ranking without consulting faculty and students. Therefore, students are still facing connectivity issues and have serious concerns about the quality of online education and summative assessment. One major step taken by the Government is to promote all the students in schools and colleges in the next grade without examination. However, the assessment of university students is still lurking.

Ashraf (2020) reported that one of the leading newspapers of Pakistan The News has surveyed to find opinions of both faculty and students about online education in this pandemic. Most of the students are not satisfied and facing many problems, i.e., connectivity and overburden online lectures. Faculty is also not happy with this new scenario where students are not under their direct influence or contact of teachers. Faculty and management consider that recorded lectures are more feasible and productive because the student can play it according to his/her convenience. Students have urged the government and higher educational institutes to provide subsidized loans to the students so that they can buy laptops for online education.

Sikandar (2020) A meeting of think-tank is Allama Iqbal Open University unanimously conceded this crisis as an opportunity for higher education institutes to enhance the quality of online education. HEC and the government should play a vital role in eliminating the digital divide, especially in remote areas like Baluchistan. The orthodox instruction methods now necessitate being substituted with “equip and empower students” approach. Broad cooperation between public and private sector universities is indispensable for superior and swift solutions. The education advisory council should be established to bridge gaps between theory and practice.

Above mentioned all these studies cogitated on the online education and have addressed the new modifications and challenges in uncertain circumstances of education under this pandemic emergency. Therefore, the word ‘online education’ is being focused on comprehending the keyness, collocations, and visualization of the particular corpus (constructed on online education in Pakistan exclusively retrieved through electronic media) in the context of COVID 19. The present research will unfold a better understanding of the readiness, challenges, and effects of the pandemic on the higher education sector of Pakistan.

IV. METHODOLOGY

The choice of the dataset is based on the generation of automatic web-based corpus through scraping the websites of four leading English newspapers of Pakistan. An extensive collection of texts used for studying language is called corpus. In corpus analysis, words are labeled with parts of speech (POS) and grammatical description. The terms corpus and text corpus and language corpus are interchangeable. Using a corpus for any linguistic or language-oriented work ensures the outcomes reflect the real use of the language. There were multiple goals to generate a web-based corpus of publications electronic media. Firstly, the tool, sketch engine was used for web scrapping; the relevant data was fetched through providing seed words. Therefore, there was considerable convenience to include the comments, responses, and feedback besides the content of newspaper articles. Secondly, to construct a large corpus based on a wide range of topics related to online education and maximum participants with different social, economic, and political backgrounds. Thirdly, to examine utmost possibilities, for example, Word Sketch, Thesaurus, Concordances, Keywords and N-grams (Multiword expressions or MWEs), finding out the themes and values of the corpus and further its comparison to survey results. Finally, data was analyzed quantitatively and objectively. The selected newspapers are The Dawn, The Daily Times, The News International, and The Express Tribune. A corpus detail is provided through a table below. These values were found after cleaning the data by using Wordsmith tool version 1.8. It provides separately, type and token and keywords by running the corpus init.

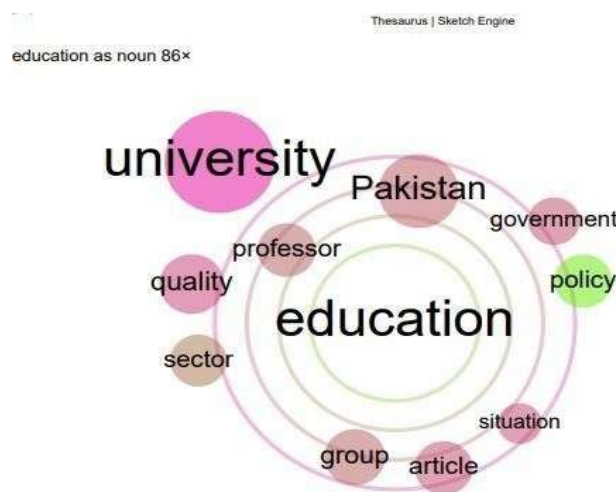
Table. 1
Corpus and Sub-corpus Values Retrieved through Wordsmith Tool 1.8

No.	Newspapers (URLs)	Total Words	Type	Token	Keywords
1.	The Dawn https://epaper.dawn.com	7,242	1,180	7,804	59

In Fig. 2, an automatically produced thesaurus identifies words that occur in similar contexts as the target word 'education'. It provides the information of distributional semantic categories. As the automatically produced thesaurus is available for each word in the corpus, but currently, the word 'education' is focused. The red text area is the result that matches the search criteria, for example, 'university', 'quality', 'policy', 'situation', and 'professor'.

N-gram is a sequence of several items (bigram = 2 items, trigram = 3 items ...n-gram = n items)? An item can refer to anything (letter, digit, syllable, token, word, or others). In the context of corpora and corpus linguistics, ngrams typically refer to tokens (or words). In linguistics, ngrams are sometimes referred to as MWEs, i.e., multiword expressions. (www.sketchengine.eu/guide/glossary/) Generating a list of the most frequent n-grams based on 4-gram helped to see the linguistic phenomena in the current corpus. According to the results of the Sketch Engine, it is found as 'the quality of education' most frequent 4-gram string of the corpus. The computational technique of Ngrams can identify discourse markers or chunks of language. It showed that online education should be taught/learned with more emphasis.

Fig. 2 Sketch Engine Results of the “Thesaurus Component”



It is clearly visible that education, university, government, policy, and quality are the most frequently used terms in this corpus. The detailed concordance list, keywords list, and is attached as Appendix B and C.

VI. SURVEY RESULTS

The second part of the research design was the comparison of text analysis and survey results regarding the outcome of emergency online system. At the University of Education, Lahore, one of the largest public-sector universities with multiple campuses, a survey regarding students' online class experiences was conducted from 8-11 June 2020. Overall, 7198 Students of University of Education, Lahore responded to the online survey based on 20 questions. (University of Education, Lahore, Survey of Students' Online Classes Experiences).

For the present research, finding out comparative results among two methodologies, i.e., text analysis and survey report, this survey has been referred. Overall survey results are described below as a narrative form, and figure 20, based on the last comprehensive question about online education, is presented to quote facts and figures.

In this survey students were asked to give opinion about availability of devices used for accessibility to online education, proper internet connection, course content delivery, benefits of the online class, assessment procedure, teachers, adequate and timely feedback, instruction for online assignments, engagement of online courses, video lectures, the relevance of material and fulfillment of learning needs were asked. The students remained unsatisfied with most of the above issues. They have shown their dissatisfaction with this new mode of learning, except agreeing with the cost-effective travel perspective.

Fig. 3 Survey Results for Comparative Analysis

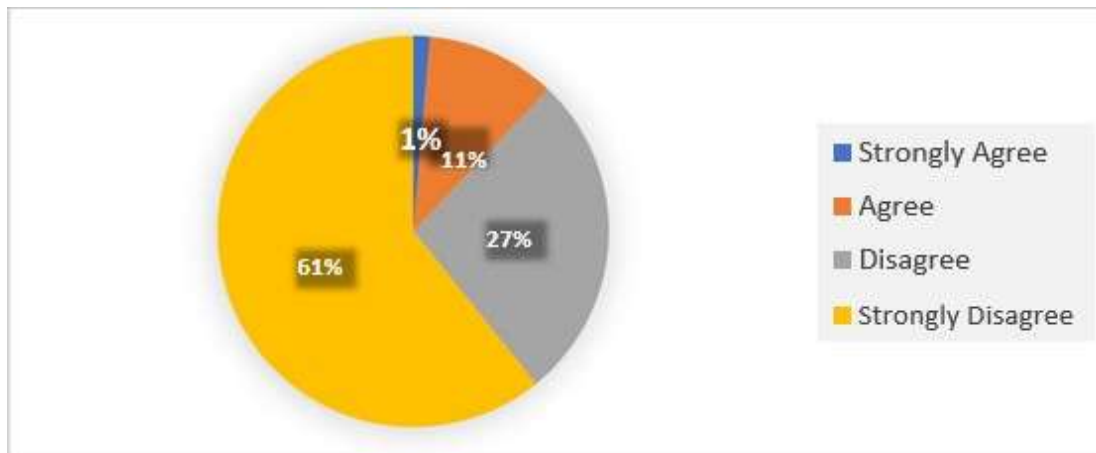


Figure 20: Overall, you are satisfied with online classes

VII. CONCLUSION

The present study is based on the keyness, collocations, and visualization of the self-constructed corpus of electronic media to explore the emergency online education system during COVID-19. After establishing the theoretical background based on newspaper writings of the experts and stockholders, it was summarized that the new change and challenge of the education system was uncertain, unpredictable, and uncontrollable. To observe the phenomenon, a comprehensive research design was built up based on the text analysis and its comparison to survey results. During text analysis, the strong collocates were found as 'policy', higher, and 'sector'. It highlights the great concern and importance of online education. The weak collocates represent the word, 'emergencies, panic, and ideas. It showed some insecurity about the new system. The relevant keywords are found as 'focal person', education policy, and education system. It seems that the purpose behind writing was grabbing the attention of authorities. The results of the thesaurus component in Fig. 2 also showed the alignment with keywords and collocation results. In survey results, the students remained unsatisfied with most of the issues related to online education. They have demonstrated their dissatisfaction with this new mode of learning. Based on the results, a few extensions in which the present research may be helpful are expressed. e. g. orientation of the new system of education is imperative through institutionalized training for teachers, students, and authorities. To create maximum awareness of the digitalized world and its most convenient accessibility. The post-pandemic measurements should evolve a hybrid system of education containing face to face and online learning simultaneously. The research will also be helpful for teachers, students, and policymakers to examine the new dynamics of learning and education. Finally, it is expected that the current research will strengthen the claim that the online education system is a newly born phenomenon that will grow through extraordinary provisions and responsiveness. Consequently, this new changes and challenges stimulate further investigation.

REFERENCES

1. Abid, A., Jabbar, A., Saeed, K., Akhtar, N., and Iqbal, J. 2013. A descriptive study: integration of technology bring flexibility to distance education in Pakistan. *Journal of Education and Practice* 4(13):111-118.
2. Ahsan, F. 2018. List of universities that provide online education in Pakistan. 13 December
3. <https://www.whateverison.com/list-of-universities-that-provide-online-education-in-pakistan/>
4. Alvi, U. 2020. *Online Education and University of Education*. The Daily Times, 16 May <https://dailytimes.com.pk/613115/online-education-and-university-of-education/>
5. Ashraf, M. 2020. *Online Learning: A Blessing in A Change World*. The Express Tribune, 06 May <https://tribune.com.pk/story/2214874/6-online-learning-blessing-changed-world/>
6. Abbas, F., Jalil, M. K. & Rehman, Z. (2019). The Status of Punjabi language in the province of Punjab, Pakistan. *The Dialogue*, 14 (2), 255-264.
7. Abbas, F., Farid, M. F., Iqbal, A. & Parveen, S. (2020) Impact of using newspapers reading on English

reading proficiency: A study of Pakistani university students. *International Journal of Innovation, Creativity and Change*. 14 (10), 223-232.

8. Banuri, T. 2020. *Why Online Education*. The Dawn, 13 April https://epaper.dawn.com/DetailImage.php?StoryImage=03_04_2020_006_004
9. Bari, F. 2020. *Online Learning*. The Dawn, 03 April
10. Batool, S. 2020. *COVID-19 Impact: Digital Learning Lays Bare Economic Disparities*. The Express Tribune, 30 May <https://tribune.com.pk/story/2210890/8-covid-19-impact-digital-learning-lays-bare-economic-disparities/>
11. Burhan, K. 2020. *Coronavirus Stunts, learning as Pakistan's Education System Is Thrown into Disarray*. The News, 11 June
12. <https://www.thenews.com.pk/latest/640402-coronavirus-stunts-learning-as-pakistans-education-system-is-thrown-into-disarray>
13. Bhatti, A.M., Abbas, F. & Rana, A.M.K. (2020). An Empirical study of learning styles used by undergraduate English learners in public sector colleges in Pakistan. *Elementary Education Online*, 19 (3), 1864-1875.
14. Ahmed, S. N., Abbas, F. & Naz, F. (2020). Historical development of orthography in English and impact of computer-mediated communication (CMC) on the emerging orthographic patterns in English. *PalArch's Journal of Archeology of Egypt* 17(11), 162-175.
15. Hoodboy, P. 2020. *HEC- Seize the Corona Moment*. The Dawn, 25 April <https://www.dawn.com/news/1552055>
16. Kahlid, R. 2020. *Aligning Higher Education with COVID-19 Realities*. The News, 17 May
17. <https://www.thenews.com.pk/print/659839-aligning-higher-education-with-covid-19-realities>
18. Mehmood, A. 2020. *COVID-19: Education in Pakistan Gets Mundane Attention*. The News, 19 May <https://www.thenews.com.pk/print/660617-covid-19-education-in-pakistan-gets-mundane-attention>
19. Qazi, W. 2020. *Online Education Is Here to Stay*. Express Tribune, 28 April <https://tribune.com.pk/story/2208369/6-online-education-stay/>
20. Rehman, M. 2020. *Online Education in Pakistan in COVID-19 era*. The Daily Times, 30 April
21. Sikander, A. 2020. *Online Teaching and Our University*. The Daily Times, 03 May <https://dailytimes.com.pk/writer/dr-arif-sikander/>
22. Glossary Sketch Engine. <https://www.sketchengine.eu/guide/glossary/>
23. Hill, J., and M. Lewis, eds. 1997. *Dictionary of selected collocations*. Hove, UK: Language Teaching Publications.
24. McCarthy, M. (1995) *Vocabulary*. Oxford: Oxford University Press.