



ADVANCE SOCIAL SCIENCE ARCHIVE JOURNAL

Available Online: <https://assajournal.com>

Vol. 05 No. 01. Jan-March 2026. Page#4094-4106

Print ISSN: [3006-2497](#) Online ISSN: [3006-2500](#)Platform & Workflow by: [Open Journal Systems](#)<https://doi.org/10.5281/zenodo.21337500>

A Corpus-Based Sentiment Analysis of University Students' Attitudes Toward Online Learning: Evidence from a Pakistani Public University

Farhat Noor

M.Phil Scholar, Department of English, University of Okara, Okara, Pakistan.

farhatnoor89@gmail.com

Dr. Muhammad Ilyas Mahmood (Corresponding Author)

Assistant Professor, Department of English, University of Okara, Okara, Pakistan.

ilyas.mehmood@uo.edu.pk

Ali Ahmad Sial

M.Phil Scholar, Department of English, University of Okara, Okara, Pakistan.

ahmali8094@gmail.com

Ayesha Tufail

M.Phil Scholar, Department of English, University of Okara, Okara, Pakistan.

official.ayeshatufail@gmail.com

ABSTRACT

This study explores the attitudes of the university students to online learning in a public university in Pakistan through a mixed method approach, both an explicit survey based quantitative approach and an implicit corpus-based sentiment analysis approach. With the rapid spread and adoption of digital learning platforms, the higher education system in Pakistan has experienced a major transformation, and it is imperative to understand the real cognitive and emotional reactions of students to the changes in this system. A specialized Learner Corpus was created out of 23 student essays (4396 word tokens and 999 word types), and survey data from 36 university students from different disciplines (English, Computer Science, Education and Business Administration). This study uses corpus analysis tools (AntConc v4.2.1) and a Corpus-Assisted Discourse Studies (CADS) framework to systematically find out what lies beneath the surface with regard to the semantic prosodies, in the form of absolute frequency lists, collocational networks, and concordance lines. The quantitative survey results reveal a high level of agreement with respect to the theoretical advantages of digital education with high agreement scores in the areas of spatial flexibility and convenience. But the power of the corpus analysis clearly outweighs such findings, with positive sentiments having strong linguistic associations with learner autonomy, and negative sentiments being strongly tied to infrastructural decay. The words that tend to group with internet (problems, issues) and with interaction (lack, poor), create a discourse in which students are passive victims of the power outages and the "digital divide. The study concludes that although the students of Pakistan are enjoying the time saving by using digital platforms, there are significant pedagogical barriers due to poor infrastructure and lack of face to face interaction. The results strongly recommend the use of a hybrid education model by institutions, and a strategic shift to using asynchronous digital interventions.

Keywords: Corpus-Based, Sentiment Analysis, University, Attitudes, Online Learning, Evidence, Public University.

1. Introduction

1.1 Background of the Study

The study began with a background that led to the following research questions: Information and Communication Technology (ICT) has brought a paradigm shift in pedagogy in higher education on global level. The drastic transformation towards online learning, triggered by global crises, changing socio-economic needs and fast-paced modern technology advances, has created a novel learning medium, which by its very nature questions the traditional face-to-face educational model (Kumar, 2024). In the particular case of Pakistan universities, this change has been unparalleled and intricate. Public universities of Pakistan are in a unique pedagogical duality as they are located at the crossroads of a fast digitalizing youth and a growing national infrastructure (Shakeer & Husni, Talita Afifah, 2024). Synchronous and Asynchronous digital platforms, such as Zoom, Google Meet, Microsoft Teams and WhatsApp have been widely integrated and are ubiquitous across academic departments (AkinFakorede, et al., 2026). But, as is often the case with such transitions, it hasn't been uniform, seamless, or painless. Online learning in Pakistan is both dependent and marginalized by the socio-economic inequalities in access to technology and yet empowered by the democratization of technology. So, in Pakistan, higher education institutions are now moving towards a 'transitional' stage where there is a mismatch between expectations of "traditional pedagogy" and the "digital reality" (Khan et al., 2026). In a very dynamic and often highly stressed educational environment, it is all the more important to know what university students are truly feeling, thinking and doing about online learning in order to ensure a sustainable future of digital learning in this region (Fadele, & Ajibola, 2026).

1.2 The problem statement

While there have been many studies of online learning in the Pakistani context of higher education, there is a serious lack of any kind of computational/corpus-based evaluation of students' discursive attitudes. The feedback mechanisms still in use in the institutions are almost entirely based on quantitative surveys using a likert scale. These traditional surveys have the disadvantage of not being able to systematically reveal the affective unconscious language mechanisms by which students build and express their digital learning experiences. As such, there is a paucity of meticulous empirical evidence on the discursive negotiation of the theoretical advantages of the digital flexibility by students with the real-life challenges of infrastructural marginalization (the so-called 'digital divide'). This study fills this important methodological gap by analyzing the same phenomenon of sentiment analysis of online learning by comparing and contrasting the findings of a corpus-based study and quantitative survey results to reveal the underlying, linguistically-based realities of online learning in a Pakistani public university.

1.3 Research Objectives

This study aims to fill certain gaps in the existing literature in a systematic way by addressing the following main academic objectives:

1. To provide numbers for absolute frequency and distribution of important key vocabulary and semantic fields for online learning in a specialized, student-generated text corpus.
2. To cross-check the implicit, computational corpus evidence with explicit quantitative survey evidence, to offer a highly holistic and cross-validated impression of university students' attitudes towards online education.

3. To critically analyse the implications these linguistic patterns have on broader issues of ideology, pedagogy and infrastructure in the context of the “digital divide” in Pakistani Higher education.

1.4 Research Questions Based on the above research aims, there are three specific research questions in this study:

1. What is the most common lexicon and semantic area used by the university students when they naturally talk about their experience with on line learning?
2. What information about the students' underlying, subconscious sentiment and emotional orientation do the specific collocational networks and concordance patterns of highly charged target words like flexibility, accessible, lack, poor and challenges provide?
3. How do the implicit linguistic findings from the specialized learner corpus complement, challenge, and/or go beyond the explicit attitudes that are revealed in the quantitative Likert-scale survey?

1.5 Significance of the Study

This research is important and opens new avenues for the growing and very interdisciplinary discipline of Corpus-Assisted Discourse Studies (CADS) as it applies the methodological frameworks to the field of educational technology in a South Asian context. The study's approach of combining computational linguistics (CL), with the help of computer software like AntConc, with traditional quantitative survey techniques is successful in overcoming the binary opposition of qualitative and quantitative research that is often found in educational research. It offers statistically valid evidence that is linguistically nuanced to university administrators, curriculum developers, higher education policymakers and educators that informs their understanding of the realities of students' lives. Moreover, this paper goes beyond the black-or-white judgements that are often thrown at online learning, and shows how the "digital divide" is a particular issue that needs to be specifically addressed in developing countries. The results of this study provide an important basis for building highly context-abased, hybrid pedagogical models that are based on the infrastructural and socio-economic contexts of the Global South.

2. Literature Review

2.1 Global Perspectives on E-Learning and Students' Attitude

The academic pendulum has swung a long way in the last decade from the perspective that online learning is a nice to have to an understanding that it is a now to have – one that is available to all people, everywhere, every day. Online learning is viewed from a predominantly constructivist perspective in international educational discourses as very democratising, providing flexibility of space and time (Muraleedharan, 2024). Much in the literature around the world, attention is drawn to the power of Information and Communication Technology (ICT) to overcome the limitations of geographical space and provide access to world class education resources, e-libraries and AI pedagogic tools to diverse groups of learners. International surveys of student attitudes have found very positive feelings about the freedom afforded by the asynchronicity of online learning, such as, pausing, re-winding and re-watching recorded lectures and lectures generally is reported as a significant cognitive benefit. In addition, the international story has been one of a cost-effective and convenient online alternative; students save a lot of time and a lot of money because they don't have to travel physically. This is because, in high-achieving settings with high-quality digital systems, student attitude is largely positive, and, as part of the modern learning experience, students are keenly accepting of the flexible and self-directed nature of digital learning (Stutchbury et al., 2025).

But a largely positive and monolithic global perspective like this, according to postcolonial and sociological critique of education, is often built on the assumption of ubiquitous technological access and neglects the harsh infrastructural situation and systemic inequalities of the Global South.

2.2 Digital Divide in the Developing Nations

However, if one moves to the scenario of Pakistani educational system, when it comes to Online learning, the situation is much more complicated and problematic. A deep and pervasive "digital divide" in HE(higher education) has been emphasized in recent research on HE in Pakistan (Jamil, 2023). This segregation means that a student's socio-economic and geographical backgrounds largely influence their online learning experiences. Students living in a metropolitan area may have a digital education as much as their counterparts have in other parts of the world, but the students living in semi-urban and rural areas including the areas of South Punjab suffer from infrastructural weaknesses that make their education debilitating (Wood, 2023). The academic literature regularly cites the constraints of frequent load shedding, astronomical cell phone data packages, poor broadband reception, and not having a personal digital device (laptop or reliable smart phone) as having a significant effect on the experience of e-learning (Ghoulam et al., 2024). In this context, the fruits of theoretical knowledge of the advantages of digital platforms and technological advancement have proven to be very disappointing when it comes to the actual implementation of online classes and coursework as it leads to overwhelming academic dissatisfaction, cognitive overload and deep sense of isolation among students in Pakistani academia. It is argued in the literature that this situation of constant struggle with technology and lack of face-to-face interaction de facto changes the pedagogical landscape and that the student experience is no longer digital empowerment but infrastructural marginalisation and anxiety (Maposa & Blessing, 2023).

2.3 Corpus Linguistics and Sentiment Analysis in Education Studies

Modern educational researchers have increasingly resorted to the sophisticated methodologies of Sentiment Analysis and Corpus Linguistics to carry out in-depth analysis of these complex and multifaceted attitudes of students (Wang et al., 2025). Qualitative thematic analysis, despite its decidedly important value for localised case studies, is not necessarily wide or scientifically sound enough to systematically record patterns of language use in large data sets (Ahmed et al., 2025). In contrast, corpus linguistics can provide a very powerful quantitative, objective and overarching method to uncover hidden linguistic characteristics - word frequencies, statistically significant words and complex collocational networks - that are often completely beyond the scope of the naked eye (McEnery & Brookes, 2024). The method adopted is Corpus-Assisted Discourse Studies (CADS) which enables researchers to investigate the notion of "semantic prosody" (as put forward by Sinclair and Louw) that assumes that words have a positive or negative aura that resides below the level of consciousness and are the product of their frequent association with other words (Gillings et al., 2023). Thus explicit and implicit semantic prosody of a text can be objectively and mathematically measured by systematically mapping the co-occurrence of specific evaluative adjective modifier words with target words (e.g. how often is poor used with internet or lack with interaction? (Zasowska, 2026) This methodology has an important role in educational research since it exposes students' hidden ideologies and emotions and shows how students use language to create themselves as either empowered, active students or victims to the failure of technology in a digitally mediated educational context.

2.4 Research Gap

An extensive literature search was conducted and it was found that there is a clear and significant gap in the literature. The traditional survey research methods have clearly demonstrated that e-learning has a very high level of infrastructural challenges in Pakistan, and on the other hand, CADS is an exceptionally strong methodology for sentiment analysis, but there is a dearth of studies where these two fields are studied together. In particular, there is no study that has methodologically used a corpus-based sentiment analysis to analyze the naturalistic text produced by Pakistani university students when they speak about online learning; coupled with explicit quantitative survey data for triangulation. This current study helps in bridging this significant empirical gap and offers a linguistically informed and in-depth exploration of the multifaceted reality of the digital divide as experienced by the Pakistani students.

3. Methodology

3.1 Research Design

The research design for this study is a very robust mixed method, that is convergent parallel approach in order to fulfill the comprehensive research objectives as above. This methodological approach deliberately brings together quantitative (methodological) as well as qualitative (computational) corpus-driven (linguistic) survey analysis (McEnery & Brookes, 2024). The major reason for the use of mixed method is for triangulation. The study will have a high level of internal validity because it will measure both explicit, consciously expressed opinions from the students by using the Likert-scale metrics, as well as implicit, subconscious emotional and ideological orientations using the natural language that the students use in their essays. This triangulation ensures that the explicit attitudes that are reflected through the traditional survey scales are thoroughly examined and verified with the implicit attitudes that are in the students' authentic textual production.

3.2 Participants and Setting

The subjects for the present study consisted of 36 university students of a well-known public University in Pakistan with diverse characteristics. The sample for the respondents were carefully chosen in such a way that it would represent a cross section of the population within the higher education community. The participants were both male and female students and the vast majority (92%) were in the age range of 20 to 29 years, which is the typical age range of modern digital natives in higher education. In addition, participants came from various levels of academic study, both undergraduate (Bachelor of Science - BS), and post graduate (Master of Philosophy/Master of Science - MPhil/MS). To ensure a diversity of disciplines and prevent faculty-specific biases on technology, the respondents were drawn from a number of different academic departments: English Language and Literature, Computer Science, Education and Business Administration. This diversity in sampling makes it highly likely that the results would not be representative of the views of a small group of very specialized academics, but of the general attitude of the institution.

Table 1: Demographic Profile of Survey Respondents (\$N=36\$)

Demographic Category	Classification	Count (n)	Percentage (%)
Gender	Male	18	50.0%
	Female	18	50.0%
Age Group	20–24 years	21	58.3%
	25–29 years	10	27.8%
	30 years or above	5	13.9%
Degree Program	Bachelor of Science (BS)	17	47.2%
	Master's (MPhil/MS)	19	52.8%

Academic Department	English Language & Literature	19	52.8%
	Computer Science & IT	4	11.1%
	Business Administration	2	5.5%
	Education	2	5.5%
	Other / Mixed Disciplines	9	25.1%

3.3 Data Collection Procedures

To fulfill the mixed-methods design, there were two separate yet complementary procedures for data collection. Survey Data: Quantitative data was gathered using a very structured and researcher designed survey, which was electronically sent out using Google Forms. In order to quantitatively examine the students' attitude in relation to several major thematic areas, a standardized 5-point Likert scale (between Strongly Disagree and Strongly Agree) was used. These domains encompassed perceived flexibility of online education, development of digital skill, poor connectivity and technical problem, lack of pedagogical interaction and students' preference for future direction of higher education (e.g., preference to hybrid education model). The qualitative and linguistic data was collected by the creation of a specialised digital Learner Corpus. This corpus was created by collecting 23 short essays and reflective academic paragraphs from the surveyed learners where they were asked to describe their personal experiences, problems and gains related to online learning. This data was then manually pre-processed in detail and to a large extent, to be ready for computational analysis. All metadata, respondent IDs and irrelevant structural formatting were completely stripped. The text was cleaned up by being consistently punctuated, by being corrected for spelling deviations that affect the proper tokenization of it, and by being converted into clean, machine-readable plain text, using UTF-8 encoding. Finally, a very specialized final corpus was created after this thorough cleaning process: 4,396 word tokens and 999 unique word types.

3.4. Analytical tools and framework

Data collected was analysed using two analytical approaches which reflect the nature of the data. The native analysis features of Google Forms and Microsoft Excel were used to compute strong descriptive statistics such as percentage distributions, frequency counts and central tendency measures for all items on the Likert-scale for the quantitative data obtained from the survey. The industry standard desktop computer corpus analysis software AntConc (version 4.2.1, developed by Laurence Anthony) was used on the specialised Learner Corpus. AntConc was used to systematically retrieve absolute word frequencies, to create ranked lists of key words and to visualize large collocational networks. In particular, a targeted collocation and concordance line analysis was carefully carried out on predefined node words having a positive sentiment (like effective, flexibility, convenience and accessible) as well as on node words with a negative sentiment (like challenges, problems, issues, lack and poor). This computational framework enabled the researchers to show decisively how certain lexical collocates discursively contribute to the overall stories of technological empowerment or systemic infrastructural marginalisation.

4. Data Analysis and Results

4.1 Quantitative Survey Findings

The quantitative data collected from a structured survey (N=36) that used the Likert scale, provides a very telling statistical analysis of explicit attitudes of the Pakistani university students towards online learning. The data points to a deeply split educational experience, and one that is deeply conditioned by the tension between theoretical conceptualization of pedagogical convenience, and stark infrastructural realities. The data shows that there is a

strong dependence on the few dominant digital ecosystems when it comes to analysing platform usage. Zoom, Google Meet and WhatsApp were the most popular platforms to use for teaching and student communication, while Microsoft Teams and Google Classroom were used less. This dependence on synchronous video platforms immediately paves the way towards the infrastructural conflicts mentioned in the following section of the survey.

When it comes to spatial and temporal autonomy, the survey shows an amazing uniformity of opinion as to the advantages of digital education. The majority of the respondents chose "Agree" or "Strongly Agree" when answering the statements "Online learning provides flexibility in managing study time" and "I find online learning convenient for academic activities. The two statements highlighted the main advantage of the online modality: getting rid of the time and spatial limitations of the traditional physical campus.

Table 2: Key Quantitative Survey Findings on Student Attitudes (Likert Scale)

Survey Statement / Prompt	Strongly Agree / Agree (%)	Neutral (%)	Disagree / Strongly Disagree (%)
Online learning provides flexibility in managing study time.	72.2%	16.7%	11.1%
I find online learning convenient for academic activities.	69.4%	13.9%	16.7%
Poor internet connectivity negatively affects online learning.	94.4%	2.8%	2.8%
Online classes reduce interaction with teachers.	83.3%	11.1%	5.6%
It is difficult to remain focused during online classes.	77.8%	11.1%	11.1%
Technical problems frequently disrupt online learning.	88.9%	5.5%	5.6%
A combination of online and face-to-face learning is preferable.	88.9%	8.3%	2.8%

But this clear recognition of flexibility is quickly and overpoweringly undermined by responses on issues of infrastructural barriers. The overwhelming majority of the participants (99%) agreed that poor Internet connectivity has a negative impact on online learning when given the statement: Poor internet connectivity negatively affects online learning. This is the most negative data point in the whole survey, and it shows connectivity is not just a minor inconvenience: it's a major impediment to learning.

Moreover, the survey revealed great social and psychological gaps related to digital transition. The results to the questions 'Online classes decrease the interaction with teachers' and 'It's hard to stay focused in online classes' were strongly in favor of the statement, indicating a clear lack of a community learning environment. In conclusion, students' vision for the future was very pragmatic, despite the many infrastructural and psychological obstacles. A near-unanimous response was given for the statement "A combination of online and face-to-face learning is preferable", indicating that students do not want to give up digital learning tools and they would like to have a hybrid future for learning, but they also want to have improvements to the institutions for online learning, as reflected in the high agreement with the statement "Universities should improve online learning facilities.

4.2 Corpus Frequency and Keyword Analysis

The computational analysis of the Learner Corpus (4,396 word tokens, 999 unique word types) yields key empirical data and information about the most prominent thematic and cognitive spaces in which the students are located, whereas the quantitative survey invites the students to express their conscious and explicit opinions. Through the extraction of absolute frequencies with AntConc (v4.2.1), we can objectively determine the key words used as the core lexical items which frame the students' digital learning discourse.

Table 3: High-Frequency Lexical Items in the Specialized Learner Corpus (Total Tokens = 4,396)

Rank	Lexical Item	Absolute Frequency	Primary Grammatical/Semantic Role
1	online	137	Modifier / Core Thematic Node
2	students	130	Primary Social Actor / Experiencer
3	learning	107	Primary Pedagogical Process
4	classes	48	Contextual Spatial/Digital Setting
5	time	29	Temporal Resource / Constraint
6	education	28	Broad Institutional Context
7	internet	24	Primary Infrastructural Dependency
8	problems	19	Primary Negative Sentiment Marker
9	technology	17	Medium of Pedagogical Instruction
10	lack	14	Negative Deficit / Absence Marker
11	flexibility	13	Primary Positive Sentiment Marker
12	interaction	13	Pedagogical / Social Deficit Marker

A study of this frequency distribution shows that there is a definite cognitive emphasis on the part of the respondents. Core thematic words like online (137 occurrences), students (130 occurrences) and learning (107 occurrences) unsurprisingly head the absolute frequency list, but the next few words are very representative of the students' lived experiences. The word 'time' appears 29 times, indicating the high importance of the cognitive aspect of temporal management as a resource that is saved (e.g., commute time) and a constraint. One of the most striking things is that words that show the systemic barriers and deficits are mentioned with startling regularity—such as internet (24), problems (19) and lack (14)—and outnumber the positive words, such as flexibility (13). The resulting lexical hierarchy shows that the students' consciousness is very much focused not only on learning itself, but also on the ubiquitous and on-going technological barriers which mediate and often impede learning.

4.3 Concordance and Collocation Analysis: Positive Sentiments

The students should be aware of positive sentiments which are expressed in the passage. To systematically break down the ideology and emotion that surrounds the corpus frequencies, a specific keyword node with positive sentiment was focussed on throughout a targeted concordance and collocation analysis. The findings of the analysis of terms like 'flexibility' (13 occurrences), 'effective' (11 occurrences), 'accessible' (5 occurrences) and 'convenience' (5 occurrences) show a very consistent semantic prosody associated with the liberation of space, effective use of resources and the democratisation of knowledge.

Table 4: Frequency and Semantic Prosody of Target Sentiment Nodes in the Learner Corpus

Target Word	Node	Absolute Frequency	Polarity	Dominant Semantic Prosody / Contextual Usage
flexibility		13	Positive	Temporal autonomy, spatial liberation, managing responsibilities.
effective		11	Positive	Outcome-based evaluation, pragmatic

			utility of digital resources.
accessible	5	Positive	Democratization of knowledge, breaking physical constraints.
convenience	5	Positive	Ease of access to recorded lectures and materials.
internet	24	Negative	Site of infrastructural struggle; heavily collocated with "poor."
problems	19	Negative	Technical/academic disruption; linked to power and connectivity.
lack	14	Negative	Deprivation of human connection; collocated with "interaction."
issues	12	Negative	Systemic unreliability disrupting the learning process.
poor	11	Negative	Evaluative modifier exclusively targeting infrastructural quality.
challenges	10	Negative	Obstacles to engagement, focus, and equitable learning.

The analysis of the concordances for node *flexibilitat* shows that it is always associated with verbs of provision and appreciation. In terms of language, such concordances show statements like, "many students appreciate the flexibilities of online learning," and "it provides flexibility and allows students to continue their studies. The grammatical form of the positive terminology tends to put the student in a position in which he or she is an active subject who can move and change his or her educational environment. For example, the corpus underlines spatial and physical liberation with regard to the word 'accessible': "allows students to attend classes from anywhere making education more accessible" and "makes education more accessible for students with disabilities.

Likewise, the node *effective* is used in assessing the outcomes and in justifying the use of digital tools. Statements like 'A positive attitude is essential for effective online learning' or 'Online learning can be a more effective and enjoyable learning experience' suggest a very pragmatist, outcome-focused attitude. These positive language contexts position the digital platform as the subordinate one, and the "Actor" (the student) is the one who is actively using it and getting unprecedented autonomy to perform their academic, personal and professional tasks.

4.4 Concordance and Collocation Analysis: Negative Sentiments

On the other hand, the negative attitude that can be inferred from the corpus is dominated by terminology of infrastructural degradation, system unreliability, and high level of pedagogical isolation. A negative node analysis of the individual terms that are targets of this change (problems (19), lack (14), issues (12), poor (11)) shows a clear grammatical and semantic change. Such settings are unilingual in nature and the student does not have much choice but to be a victim of the failure of external technologies.

Almost all the uses of the adjective *poor* are pre-modifying infrastructural terms, creating a very strong negative semantic prosody on the notion of digital connectivity. In these transitivity structures, the internet, the power grid, and other technology becomes the agent, the student and the learning process becomes the passive "Goal". e.g. "poor internet connections and power outages can interrupt the learning process" or "Technical issues like poor internet connectivity, especially in rural areas, disrupt learning". The term *internet* (24 instances) is not

a very empowering one: it is a site of constant struggle and is tightly associated with collocates such as unstable, problem and issues.

Moreover, the noun 'lack' reveals a deep psychological and social impact of the Digital Transition. It is powerfully and repeatedly associated with words that denote "face to face" contact, for instance in the very common collocation "lack of face-to-face interaction" (as in "Often, lack of face-to-face interaction leads to feelings of isolation and low motivation"). All of these negative nodes are modified by a series of negative external factors such as technical problems, electricity problems and even physiological detriments such as eye problems and other health problems, creating a frustrated atmosphere that students feel is always undermining them and that they cannot control their schooling by themselves.

5. Discussion

5.1 The Concept of Agency, Flexibility and Learner Autonomy

The quantitative survey data and the Learner Corpus have been compared and contrasted in a critical discourse analysis which shows that the linguistic image of online learning is bound and tied up with the notions of learner autonomy, socio-economic class and geographical infrastructure. In understanding the positive results it is clear that the students create a strong narrative of empowerment in both time and space discursively. The statistically high level of agreement in the survey with regard to flexibility is reflected perfectly in the active and agentive language that is found in the corpus. The respondents' answers use Hallidayan "Actor" transitivity structures to discursively "take ownership" of their educational timelines: I could go to lectures from home; students saved time on their commute; they learned at their own pace. The digital classroom is a very liberating mechanism for a demographic that have multiple lives – such as part-time work and family responsibilities. It is an effective way to overcome the physical and economical obstacles for higher education, allowing introverted students, students with mobility difficulties or staying far from the physical campus to interact with the academic material in their own pace and manner. The linguistic positioning is very strong overall in terms of an emergent sense of academic agency.

5.2 Infrastructural Marginalization and the Digital Divide

At the same time, the data reveals an unmissable materiality of infrastructure that is a serious impediment to, and fissure in, this theoretical empowerment. The strong negative semantic prosody of technology in the corpus – in particular the highly statistically significant collocations of "poor + internet", "electricity + shortage", and "technical + problems" – accurately depicts a socio-economic reality, specifically from a Global South perspective and in particular from a Pakistani perspective. The internet can be called a bridge, as is often depicted in the literature throughout the world, but for these students, the internet is often a ruthless gatekeeper in the Pakistani context. Constantly mentioned issues of interrupted power supply (load-shedding) and intermittent broadband connections symbolize inequalities that are inherent and at the core of e-learning's utopian potential. The marginalization of students from remote and rural areas (as in the specific reference to 'towns' like Vehari, Dipalpur in the raw essay data) is linguistically effected by effectively silencing them, not for lack of intelligence, but for lack of bandwidth. In such cases, the technology is no longer an empowering tool but turns into an oppressive, unpredictable entity that makes the student passive and helpless and further distances the already existing gap between urban elites and rural working class students—the digital divide.

5.3 Pedagogical isolation, human interaction

Anaphora reveals that in addition to the hardware and infrastructure gaps there is a huge pedagogical, psychological and social gap in the fully online model of education. The frequent

association of the words "lack" and "interaction" in the language of the interaction tends to suggest the basic human components that are completely absent from the digital interaction. The psychological effects in the data are very dire and students often suffer from "Zoom fatigue," extreme loneliness and a de-engaging sense of academic motivation from spending hours on a screen. The digital classroom, as elaborated in student's essays, is a place that is flat, cold, two dimensional, and lacks any of the spontaneity and energy of social interactions that facilitates meaningful and deep learning and discussion. The lack of ability to naturally pick up non-verbal cues from instructors, awkwardness of group discussions in virtual space, and physical distance from peers decreases the over-all quality of knowledge transfer. Moreover, there are some alarming concerns about academic integrity: there are even explicit references to the "risk of cheating" when speaking of online exams that are not supervised. This indicates that the virtual environment not only separates the student emotionally, but also has a tremendous impact on the authority of the institution, academic discipline and the community spirit of the university experience.

5.4 Triangulating Linguistic Sentiment with Explicit Survey Data

The explicit quantitative information derived from the Learner Corpus is combined with the implicit linguistic information from the Learner Corpus, a highly coherent and perfectly validated picture of the Pakistani student experience emerges. The language patterns identified in the corpus are naturalistic rather than explicitly provided in the survey answers, but do not conflict with the answers provided in the survey; they greatly enrich and reflect the explicit answers. The frustration and passiveness of the grammatical structures around the word internet in the essays vividly express the almost unanimous opinion in the survey that the internet has a negative effect. Likewise, for the high ratings the survey received on the flexibility, the agentive positive semantic prosody of the students' written responses to the survey can be verified. This triangulation is at the heart of the most future-oriented result in the survey: a strong majority in favor of a hybrid education. While students don't want to leave the spatial freedom and AI-powered tools (like ChatGPT and Claude) of the digital world, their linguistic dismay demonstrates that they cannot continue to live on an unsteady and isolated digital island. This is therefore a logical and direct response to the extreme duality of the educational context: a call for a hybrid model.

6. Recommendations

6.1 Hybrid/Blended Models of Institution Adoption

Based on the results of the survey and the corpus, it is clear that the participants are all in favor of getting rid of the binary all-online/all-physical approach and implementing strong hybrid (blended) pedagogical practices in higher education institutions in Pakistan. Universities should make a sensible division of their courses and their traditional teaching method: The theoretical lectures and teaching of fundamental knowledge can be recorded and broadcast online, so as to give the students the much-sought-after free time and space. On the other hand, critical pedagogical components, which cannot be conducted online without the deep human interaction/dialogue, such as complex laboratory work, group projects, academic debates and summative assessments must be mandate driven to be done on campus. This blended method is an effective way of making use of the convenience of digital technology and at the same time of directly counteracting the extreme psychological isolation and lack of interaction mentioned in the corpus.

6.2 Use of infrastructures and Asynchronous Interventions

Conversely, universities need to fundamentally re-think their approach to digital delivery, in order to directly address the universal infrastructural issues of 'poor internet' and frequent

'electricity outages'. Institutions need to make drastic efforts to decrease the dependence on platforms with high bandwidth and synchronous live streaming, since they are the ones that have a high price and have a negative impact on students from rural areas or areas where there is load shedding (like Zoom and Google Meet). Rather, teachers need to focus on learning interventions that can be done asynchronously. Universities can provide highly compressed, downloadable video lectures, reading materials that can be accessed offline, and have flexible submission windows to provide a more even-handed digital environment to prevent a student from being academically hurt when the state's infrastructure fails.

6.3 Pedagogical training and integration of AI.

Last but not least, there is a need to significantly invest in the ongoing pedagogical upskilling of the universities' teachers as well as in digital literacy of students. Teachers need to be trained formally to go beyond lecturing in a webcam and need to have strategies for how to keep students active in the online learning environment and how to use gamification and interactive digital learning tools to reduce "zoom fatigue." In addition, the use of AI tools such as ChatGPT, Claude and SciSpace are already a reality in the life of students, as seen in several examples in the corpus. However, universities must accept and embrace these tools, not simply endure them or police their use, but design in a way that enables students to learn how to use powerful platforms in an ethical, critical, and effective manner to improve their independent study skills.

7. Conclusion

This study was extensive, involving a proper corpus-based sentiment analysis along with quantitative survey evaluation to find out the true sentiments of university students towards online learning in a public university of Pakistan. The study has employed a specially built digital Learner Corpus consisting of 4,396 words and structured Likert-scale measurements, which have shed light on an extremely complex and precisely dualist attitude towards digital high education in the Global South. The quantitative and qualitative/linguistic findings clearly show that the students of Pakistan have a very positive and agentive attitude to the spatial flexibility, cost-effectiveness, and democratizing potential that online platforms have in theory.

This study however, firmly shows that this theoretical freedom is severely and constantly limited by the dismal and practical reality of infrastructural decay on a systemic level. The strong negative linguistic connotations that exist around the connectivity, power failures and harsh psychological impact of pedagogical isolation show that the current "implementation" of online learning is inequitable. In sum, this research finds that the online learning is not a panacea to replace the traditional learning system in present day Pakistanis University students and is not a pedagogical failure; it is very powerful but at the same time limited. A sustainable future of higher education (and for the Global South more broadly) requires recognizing the facts of this digital divide, discarding simply synchronous online approaches and actively creating context-specific, hybrid infrastructures that combine digital flexibility with irreplaceable, equitable, human contact.

References

- Ahmed, S. K., Mohammed, R. A., Nashwan, A. J., Ibrahim, R. H., Abdalla, A. Q., Ameen, B. M. M., & Khidhir, R. M. (2025). Using Thematic Analysis in Qualitative Research. *Journal of Medicine, Surgery, and Public Health*, 6(6), 100198. ScienceDirect. <https://doi.org/10.1016/j.glmedi.2025.100198>
- AkinFakorede, O. O., Ajani, Y. A., Malik, F. O., Adamu, J. I., Amaechi, N. M., Ahmed, A. O., & Echezona, R. I. (2026). Optimizing mobile technology for Elearning: Implications for library and information science schools in Nigeria. *Journal of Electronic Resources Librarianship*, 1–16.

- Fadele, & Ajibola, I. G. (2026). *Digital Pedagogy: Bridging the Gap Between Software and the Humanities* (Vol. 85). Springer Nature.
- Ghoulam, K., Bouikhalene, B., Babori, A., & Falih, N. (2024). Exploring the impact of mobile devices in electronics elearning: A case study evaluating the effectiveness of mobile learning applications in the field of electronics and sensors. *Advances in Mobile Learning Educational Research*, 4(2), 1058–1072.
- Gillings, M., Mautner, G., & Baker, P. (2023). *Corpusassisted discourse studies*. Cambridge University Press.
- Jamil, S. (2023). Evolving newsrooms and the second level of digital divide: Implications for journalistic practice in Pakistan. *Journalism Practice*, 17(9), 1864–1881.
- Khan, S., Ahmad, E. M., & Ahmad, W. (2026). Epistemological Continuity and Pedagogical Reform: Operationalising ProjectBased Learning in Pakistani Madrasah. *BITARA International Journal of Civilizational Studies and Human Sciences (EISSN: 26009080)*, 9(1), 230–245.
- Kumar, S. (2024). The Effects of Information and Communication Technology (ICT) on Pedagogy and Student Learning Outcome in Higher Education. *EAI Endorsed Transactions on Scalable Information Systems*, 11(2).
- Maposa, A. M., & Blessing, C. V. (2023). Managing change from faceto face to online teaching and learning in higher education during the COVID19 pandemic: Implications for equity and epistemic justice. In *Online Teaching and Learning in the COVID19 Era: Perspectives on Equity and Epistemic Justice* (pp. 157–204). Springer.
- McEnery, T., & Brookes, G. (2024). Corpus linguistics and the social sciences. *Corpus Linguistics and Linguistic Theory*, 20(3), 591–613.
- Muraleedharan, M. (2024). Does Real Learning Happen?: Flexible Online Education Through Contextualized Hybrid Model Leading to Undisrupted Constructivist Experiential Learning in a PostPandemic global Environment. In *Digitalization of higher education* (pp. 181–196). Apple Academic Press.
- Shakeer, M. M., & Husni, Talita Afifah. (2024). Transformation of Educational Information Technology: An Analysis of the Development of Islamic Universities in Pakistan. *Idarah (Jurnal Pendidikan Dan Kependidikan)*, 8(2), 127–144.
- Stutchbury, K., Ebubedike, M., Amos, S., & Chamberlain, L. (2025). Professional development in the digital age: supporting improvements in teacher education through MOOCs. *Open Learning: The Journal of Open, Distance and ELearning*, 40(1), 67–90.
- Wang, Z., Huang, D., Cui, J., Zhang, X., Ho, S., & Cambria, E. (2025). A review of Chinese sentiment analysis: subjects, methods, and trends. *Artificial Intelligence Review*, 58, 1.
- Wood, R. M. (2023). A review of education differences in urban and rural areas. *International Research Journal of Educational Research*, 14(2), 1–3.
- Zasowska, M. (2026). Semantic Prosody in Academic Book Reviews. *Engaged Languages and Literatures/Langues et Littératures Engagées: Critical Views/Regards Critiques*, 20, 187.