Situational Characteristics of Pakistani University Student English: Variation across Registers, Sub-registers, and Disciplines

Nouman Hamid[1]
Ayaz Afsar[2]

Abstract

In Pakistan, universities are one of the major domains where English is commonly used since medium of instruction in most of the universities in Pakistan is English. In universities, students are required to produce a variety of written and spoken texts. Previous studies of university student English investigated only spoken or only written registers and none of the previous studies compared academic English produced by university students in both spoken and written modes. In this context, the present study sought to explore situational variation in graduate students' spoken and written academic English in Pakistani universities across four variables: registers, registers across disciplines, sub-registers, and sub-registers across disciplines. The study is based on Pakistani University Student English (PaUSE) corpus which contains 195 in-class presentations and 329 assignments produced by graduate students from four academic disciplines (botany, education, linguistics, and management sciences) pursuing graduate degrees in universities located in five major regions of Pakistan. Based on their communicative purposes, these presentations and assignments included in the PaUSE corpus were classified into six sub-registers (two spoken and four written): general presentations, research-based presentations, article reviews, literature reviews, research proposals, and research reports. Pakistani university student English was compared on situational characteristics across four variables. For this purpose, a framework was first developed, and the texts included in the PaUSE corpus were then compared for their situational characteristics across four variables. The findings reveal that there is considerable variation in the situational characteristics of Pakistani university student English across all variables. The findings of the present study make a strong case for

---

[1] PhD Scholar, Department of English, IIU, Islamabad

[2] Professor, Department of English, IIU, Islamabad

exploring linguistic variation in Pakistani university student English that will be presented in future studies.

1. Introduction

Despite being recognized as a distinct variety of English (Kachru, 1983) in the outer or extended circle (Kachru, 1988), English in Pakistan or Pakistani variety of English did not receive the attention of linguists until early 1990s. The earlier studies of Pakistani English were concerned with a) the morphological, syntactic, and semantic changes that English underwent due to its contact with indigenous languages of Pakistan or because of Islamic traditions (Baumgardner, 1990, 1998; Baumgardner, Kennedy, & Shamim, 1993; Mahboob, 2009; Mahboob & Ahmar, 2004), b) exploring differences between Pakistani, British, and/or American varieties of English on different levels (i.e., phonological, lexical, syntactic, and semantic) (Baumgardner, 1993; Kennedy, 1993; Rahman, 1990/2010; Talaat, 1993), and c) the attitudes of Pakistanis towards Pakistani English (Baumgardner, 1995; Jabeen, Mahmood, & Rasheed, 2011). All these studies were based on subjective as well as very small data.

However, towards the end of the first decade of the twenty-first century, researchers working on Pakistani English started establishing their claims on the basis of quantitative and empirical data, often employing corpus or corpora as data source. The focus of research also shifted towards a) describing the distinctive features of Pakistani English, b) empirically testing the claims made by earlier linguists about the characteristic features of Pakistani English, and c) comparing some linguistic features in Pakistani, British, and American varieties of English using corpus of each variety (Afsar & Kamran, 2011; Jabeen, Rai, & Arif, 2011; Malik & Saeed, 2015; Mahmood, A., 2009; Mahmood, R., 2009). Although these studies established their claims on the basis of frequency information, the corpora of Pakistani English used as data source in these studies included texts only from written registers and did not include any oral language. Moreover, these corpora were not balanced as they contained a large number of texts from some registers and very few from others. Finally, these studies also investigated only a few features of Pakistani English.

During the last few years (i.e., from 2013 onwards), a growing interest in analyzing the use of English in a specific domain (register) in Pakistan also started emerging. The focus of this line of research is on describing the linguistic variation within a specific domain or register and /

or comparing the linguistic variation found in a specific register of Pakistani English with that of a comparable register of British English or American English. These studies investigated linguistic variation in written registers of Pakistani English including advertisements (Shakir, 2013), written registers of Pakistani English (Hussain, Mahmood, & Azher, 2016; Mahmood & Hussain, 2016), legal English (Asghar, Mahmood, & Asghar, 2018a, 2018b), book blurbs (Qasim & Shakir, 2016), press editorials (Alvi, Mehmood, & Rasool, 2016), press reportage (Ahmad & Mahmood, 2015), newspapers (Ahmad & Mahmood, 2016), fiction (Ali & Ahmad, 2016, 2017, 2019; Ali & Shakir, 2016), interactive registers (Shakir & Deuber, 2018), research articles (Rashid & Mahmood, 2019; Rashid, Mahmood, & Ahmad, 2017), and students' argumentative essays (Abdulaziz, 2017; Abdulaziz, Mahmood, & Azher, 2016). However, none of these studies carried out the situational analysis of the registers investigated – a very important characteristic of register analysis.

There are still other domains in Pakistan where English is widely used and there is a need to investigate the linguistic variation within those domains. In Pakistan, universities are one of the major domains where English is commonly used since medium of instruction in most of the universities in Pakistan is English. Moreover, the reading materials (textbooks, research articles, etc.) in most of the disciplines are in English. Furthermore, students have to produce a variety of written (responses to exam questions, term papers, response papers, theses, etc.) and spoken (presentations, group discussions, etc.) texts in most of the universities in Pakistan. In short, students in Pakistan are exposed to and produce a great number of written and spoken texts in university settings.

Linguists have long been interested in exploring English in academic settings, but university English has captured the attention of linguistics and researchers only in the last few decades (Biber, 2006b). However, there has been more research on registers that students are exposed to (textbooks, lectures, research articles, instructional material, etc.) (Barbieri, 2015; Biber, 2006a, 2006b; Biber & Barbieri, 2007; Biber, Conrad, & Cortes, 2004; Biber, Conrad, Reppen, Byrd, & Helt, 2002; Biber & Gray, 2016; Cao & Xiao, 2013; Carkin, 2001; Conrad, 1996b; Csomay, 2002, 2006; Egbert, 2014; Fortanet, 2004; Friginal & Mustafa, 2017; Gray, 2010, 2011; Gray & Cortes, 2011; Nesi & Basturkmen, 2006; Peacock, 2011) than the ones actually produced by them (students' term papers, assignments, presentations, etc.). In case of university students' English,

there has been more research on the written English (Adel & Erman, 2012; Chen & Baker, 2010; Conrad, 1996a, 2018; Cooper & Bikowski, 2007; Cortes, 2002, 2004; Gardner, Nesi, & Biber, 2019; Hardy, 2015; Hardy & Friginal, 2016; Hardy & Romer, 2013; Leedham & Fernandez-Parra, 2017; Staples, Egbert, Biber, & Gray, 2016; Staples, & Reppen, 2016; Tasker, 2019; Weigle & Friginal, 2015) as compared to spoken English (Csomay, 2005, 2007, 2015; Iberri-Shea, 2011; O'Boyle, 2014; Poos & Simpson, 2002; Zareva, 2009a, 2009b, 2011, 2012, 2016). Although many studies of published academic language report on the results of situational analysis, most of the studies investigating variation in registers produced by students do not include comprehensive situational analysis, which constitute a gap.

University English has not received the attention of many linguists in Pakistan. Begum (2016) is one of the studies that explored university English in Pakistan. She investigated Masters' theses and Doctoral dissertations written by university students in Pakistan by employing multidimensional analysis framework. She compared theses and dissertations with respect to discipline and section. In another study, Yousaf and Shehzad (2018) investigated the structure of 4-word lexical bundles in Pakistani university students' doctoral dissertations across three disciplines and 9 sub-disciplines. However, like previously mentioned studies on university English, these two studies are only concerned with the language produced in written mode. Moreover, neither of these two studies carried out complete situational analysis of the registers investigated, which also constitutes a research gap.

With this background, the present study seeks to investigate the situational characteristics/features of graduate students' spoken and written academic English in Pakistani universities across registers, sub-registers, and disciplines, which, in turn, will make a case for exploring linguistic variation in Pakistani university student English that will be presented in future studies. The present study seeks to answer the following main and sub-questions.

1) In what ways, if any, is graduate students' academic English in Pakistani universities similar or different on its situational characteristics?
   a. How far does graduate students' academic English in Pakistani universities vary on its situational characteristics across registers (presentations versus assignments)?

b. How far does graduate students' academic English in Pakistani universities vary on its situational characteristics across registers and academic disciplines?

c. How far does graduate students' academic English in Pakistani universities vary on its situational characteristics across sub-registers?

d. How far does graduate students' academic English in Pakistani universities vary on its situational characteristics across sub-registers and academic disciplines?

2. Methodology

2.1. Compilation of Pakistani University Student English (PaUSE) Corpus

Since there is no existing corpus of Pakistani students' academic English produced in university settings, a corpus representing students' academic English produced in Pakistani universities was compiled. The corpus was named as Pakistani University Student English (PaUSE) corpus which includes presentations and assignments produced by graduate students pursuing MS or M.Phil. degrees in four academic disciplines (botany, education, linguistics, and management sciences) in Pakistani universities. The texts for the PaUSE corpus were collected from universities located in five major regions of Pakistan, that is, The Punjab, Sindh, Khyber Pakhtunkhwa, Baluchistan, and Azad Kashmir or Pakistani administered Kashmir.

Although students in Pakistani universities produce a number of written texts, for instance, assignments, response to exam questions, theses, etc., keeping in view the previous studies, only assignments were collected as representative of students' written language (English) whereas class presentations were collected as representative of their spoken English since it is the only type of spoken text that students in Pakistani universities produce in English. Assignments are written texts that students produce outside the classroom to satisfy a course requirement with enough time for writing, editing, and revising. They may be written in response to the prompt provided by the teacher or on the basis of library research, a small-scale field research, or a proposal for a future study. These are graded and contribute to the final grade of a course. Presentations, on the other hand, consist of speech delivered formally by students in front of a class on library research or term project (Hyland, 2009). The in-class presentations are usually tied to a written assignment and are often rehearsed; therefore,

presentations are expected to share linguistic features from both written and oral registers (Iberri-Shea, 2011; Zareva, 2016). Studies reveal differences in the way L1 English students and second or foreign learners of English perceive in-class presentations. Native English students see presentations as an opportunity to present information in an informal way as well as to interact with audience and keep them involved. L2 presenters seem to be preoccupied with the informational aspect of presentations without focusing on involving their audience (Zareva, 2009a). L2 presenters are often reported to be reading aloud from PowerPoint slides or a prepared paper (Hyland, 2009). Presentations, like written assignments, are often graded and contribute to the academic progression of a student.

As far as the academic level is concerned, assignments and presentations produced by Students enrolled in only MS/MPhil programs (equivalent to eighteen years of education) were collected. This level was chosen because the students below this level lack proficiency in speaking and/or writing in English. Moreover, the academic level above MS/MPhil – PhD level - was not considered appropriate for the present study as doctoral degrees in the select disciplines are not offered in many Pakistani universities and where they are offered, the student enrollment is too low to get enough texts.

The assignments and audio-recoded presentations included in the PaUSE corpus were collected within a period of two and a half years (from Spring 2016 to Spring 2018). Once collected, the audio-recorded presentations were transcribed orthographically following conventions given in Edwards and Lampert (1993) and Biber, et al. (2004). Finally, presentations and assignments were classified into sub-registers based on their communicative purposes. For this purpose, we went through each and every assignment and presentation and labeled it with the closest possible category mentioned in the previous literature. We also sought expert opinion by asking some linguists to categorize some texts. Input was also taken from relevant class instructors in case it was found difficult to categorize a text. Six sub-registers emerged from the close examination of Pakistani university students' presentations and assignments, that is, general presentations, research-based presentations (two sub-registers of presentations), article reviews, literature reviews, research proposals, and research reports (four sub-registers of assignments). The six sub-registers found in Pakistani university student English mainly differ in their communicative purposes. Article review presents a positive and/or negative

assessment of a research article, literature review incorporates and synthesizes information from multiple sources, research proposal explains how a future problem will be investigated, and research report presents findings of an original research related to a field (Cooper & Bikowski, 2007; Romer & O'Donnell, 2011). On the other hand, general presentations are based on some library survey while research-based presentations report some original research carried out by the presenter(s).

The final corpus contains five hundred and twenty-four texts and comprises close to one million words (927,439 words to be exact). The corpus contains more assignments (329 texts, 757,764 words) than presentations (195 texts, 169,675 words), which means that 62.79% of the corpus texts consist of assignments and 37.21% are presentations.
As far as the percentage of texts (assignments and presentations combined) across disciplines is concerned, majority of the texts of the PaUSE corpus come from linguistics (35.31%), followed by 26.53% from management sciences and 23.09% from education, while the least number of texts are from botany (15.08%).

If we consider the discipline-wise breakdown of the two components of the PaUSE corpus, assignments follow the same trend as was observed for the overall corpus but with different percentages. Most of the assignments, like overall texts, are from linguistics (34.04%), followed by management sciences, education, and botany with 31.31%, 19.15%, and 15.5%, respectively. As far as the presentations are concerned, the trend is a bit different although most of the presentations still come from linguistics (37.44%), followed by education with 29.74% texts, while management sciences, and botany are represented by 18.46% and 14.36% of the presentations, respectively.

As far as sub-registers are concerned, out of two presentation sub-registers, 70.26% of the presentations included in the PaUSE corpus belong to the category of general presentations while 29.74% are research-based presentations. On the other hand, the highest number of assignments found in the PaUSE corpus belong to the sub-register of literature review (35.26%), whereas the sub-register of article review contains the least number of assignments (10.33%). Interestingly, the sub-registers of research proposal (27.35%) and research report (27.05%) are found with almost similar percentage of texts.

Finally, with respect to disciplines, general presentations have the highest frequency of occurrence in education (36.5%), followed by linguistics (24.09%), management sciences (21.17%), and botany (18.25%). Research-based presentations are predominantly found in linguistics (68.96%) with very few texts from other three disciplines. As regards the occurrence of assignment types with respect to disciplines, article reviews are mainly found in linguistics (61.76%) with a few texts from management sciences and education, whereas no assignment belonging to this category came from botany. Additionally, literature reviews are most abundantly found in management sciences (34.48%), followed by botany (25.86%), education (22.41%), and linguistics (17.24%). Research proposals, on the other hand, are found in equal percentage (38.89%) in linguistics and management sciences, while the other two disciplines have very few assignments belonging to this category. Finally, linguistics has the highest percentage of research reports (40.45%) than the other three disciplines.

2.2. A Framework for Analyzing the Situational Characteristics of Pakistani University Student English

This study adopts the register analysis framework (Biber, 1988; Biber & Conrad, 2009), which includes three components: description of the situational characteristics, analysis of pervasive linguistic features, and establishing a functional association between situational characteristics and pervasive linguistic features (Biber, 1994; Biber & Conrad, 2009). Register is taken here as a "variety associated with a particular situation of use (including particular communicative purposes)" (Biber & Conrad, 2009, p. 6). Although situational analysis and linguistic analysis are applied cyclically in a register study, situational analysis is more basic than linguistic analysis, and it is recommended that register analysis should begin with the description of the situation in which a register is produced since register, by definition, is a language variety tied to a situation and communicative purpose (Biber & Conrad, 2009). Therefore, the present study reports the results of the first component of register analysis, that is, discovering variation in the situational characteristics of Pakistani university student English across registers, sub-registers, and disciplines. This study treats presentations and assignments as registers and different types of assignments and presentations as sub-registers since they differ in their situational characteristics as well as communicative purposes.

Different frameworks of situational analysis have been used in the previous studies of register analysis; however, the most comprehensive and

detailed framework was presented by Biber & Conrad (2009) which was built on earlier frameworks presented in Biber (1988) and Biber (1994). This framework has the potential of presenting the complete situational characteristics of individual registers as well as describing the similarities and differences between or among registers (Biber, 1994). This framework helps describe the situational characteristics of registers from seven perspectives: a) "participants" (various characteristics of addressors and addressees), b) "relations among addressors and addressees" including interactiveness, c) "channel" which includes mode and specific medium, d) "production circumstances", e) "setting", f) "communicative purpose" involving both general and specific purposes along with factuality and expression of stance, and g) "topic" which includes general topical domain as well as specific topic (Biber & Conrad, 2009, p. 40).

This framework has been widely used for the situational analysis of various registers during the last few decades. However, modified versions of this framework have also been used for the description of the situational characteristics of specialized registers, especially registers within the umbrella term of "academic prose" (Conrad, 1996b; Carkin, 2001; Gray, 2011; Egbert, 2014). The motives behind modification of the framework for describing the situational characteristics of specialized registers were that a) some of the components or categories of this framework were deemed irrelevant for the description of certain registers and b) some additional categories or components were required to fully describe the situational characteristics of specialized registers.

However, none of the earlier frameworks can provide the complete description of university students' assignments and class presentations as all of these were aimed at describing the situational context of academic texts produced by experts or professionals. Therefore, we developed a framework (summarized in Table 2.1) by taking categories from some of the frameworks described in the present section. Most of the categories included in our framework are from Biber and Conrad's (2009) framework since it is the only available framework that has the potential to compare a variety of general and specialized registers produced in both spoken and written modes. However, some of the categories borrowed from Biber and Conrad's (2009) framework are treated differently in our framework. Moreover, the framework we developed to describe the situational characteristics of assignments and class presentations produced by graduate

students in Pakistani universities also includes some categories from other frameworks (e.g., Conrad, 1996b; Gray, 2011; Egbert, 2014).

The framework contains nine components or categories which are not hierarchical. The first component describes the characteristics of participants who are involved in communication event. It mainly includes the characteristics of addressor(s) and addressee(s). For addressor(s), we will describe for each text whether addressor is writer or speaker; single or group; male, female, or both (in case a text is produced by a group of students involving members of both genders). Gray (2011) and Egbert (2014) categorized the authors of journal articles, popular academic books, and university textbooks into three types based on their number: one author (single), 2-4 (small group), and five and more (large group). However, since very few texts included in the PaUSE corpus are produced by more than one author, it makes more sense to use only two categories, that is, single and group. Moreover, we will also explain whether a text is produced for single addressee or a group of addressees and whether the communication event involves the presence of on-lookers or not. Since age and first language background of the addressor(s) were not recorded during corpus compilation, these are not considered here. Moreover, some other characteristics of the addressor(s) – education, profession, etc. - are also not included since these are constant and do not vary with respect to any of the variables (register, sub-register, discipline).

The second component focusses on the relationship among participants. Here the focus will be on the possibility of interactiveness between addressor(s) and addressee(s), addressor-addressee relationship, relative status or power which will be described in terms of low-high, high-low, or equal, and amount of shared knowledge between addressor(s) and addressee(s). This component is important for describing the situational characteristics of assignments and presentations since the texts included in the PaUSE corpus are expected to vary on this component on at least one variable (i.e., mode).

Table 2.1. Framework for Analyzing the Situational Characteristics of Pakistani University Student English

I. Participants
- A. Addressor(s) (i.e., speaker or writer)
  - 1. Number of speaker or writer
    - Single
    - Group
  - 2. Gender of speaker or writer
    - Male
    - Female
    - Both
- B. Addressee(s): single or plural
- C. Are there on-lookers?

II. Relations among participants
- A. Interactiveness
- B. Addressor-addressee relationship
- C. Relative status or power: low-high, high-low, or equal
- D. Shared knowledge

III. Channel
- A. Mode: speech or writing
- B. Specific Medium: transcribed, printed, or face-to-face

IV. Production and comprehension circumstances
- A. Production: real time, planned, scripted, revised and edited
- B. Comprehension: real time, skimming, careful reading

V. Text Length
- A. text length in terms of average number of words per text
- B. Range

VI. Setting
- A. Are the time and place of communication shared by participants?
- B. Is the place of communication public or private?
- C. Is the time contemporary or historical?

VII. Communicative purpose
- A. General purpose
- B. Specific purpose(s)

VIII. Topic or Subject

IX. Evidence
- A. Level of evidence: primary, secondary
- B. Type of evidence: qualitative vs quantitative

The fifth component describes the text length. This component was not included in Biber and Conrad's (2009) framework, however, it is there in other frameworks (e.g., Conrad, 1996b; Gray, 2011). In this component, for each text included in the PaUSE corpus, we will present the mean text length as well as range. However, contrary to Conrad (1996b) and Gray (2011) who described text length in terms of average number of pages, we will use average number of words as a measure of text length. This choice is based on the reason that both Conrad (1996b) and Gray (2011) worked mainly on published texts which have standard page size, font size, line spacing, etc., the assignments in the PaUSE corpus do not seem to follow similar standards for font size or line spacing, therefore, considering average number of pages as an indicator of text length is not a reliable measure; instead, it makes much more sense to consider average number of words as a measure of text length.

The sixth component – setting – is exactly the same as is given in Biber and Conrad's (2009) framework. Here, we will describe whether the time and place are shared by participants, whether the communication took place at a public or private place, and whether the communication is related to the present or historical time. This component is also important and relevant here since it has the potential of distinguishing the texts included in the PaUSE corpus on at least one variable.

The next two components explain the communicative purpose(s) (both general purpose and specific purpose(s)) and subject or topics. Finally, in the last component, we will describe whether the texts use primary evidence or secondary evidence and in case of primary evidence, is it qualitative or quantitative. In section 3, we describe and compare the situational characteristics of Pakistani university student English with respect to different variables.

3. Results and Discussion

Biber and Conrad (2009) mention four methods for describing the situational characteristics of registers which include a) personal experience and observation, b) discussion with expert informants, c) previous research covering the kinds of registers one is interested in, and d) actual analysis of texts from the register. The situational characteristics of Pakistani university student English we present in this section is based on information gathered from all four sources. Some of the situational characteristics of Pakistani university student English were easy to discover owing to our personal

experience, observation, and familiarity as university teachers with the characteristics of students' assignments and presentations. Some other situational characteristics were identified through actual examination of the presentations and assignments included in the PaUSE corpus during transcription of presentations and cleaning up of assignments. However, in case a feature remained ambiguous despite our familiarity with the registers as well as analysis of the actual texts, help was sought either from language teaching experts or previous literature reporting the situational characteristics of similar registers. In the remainder of this section, we will describe the situational features of the texts included in the PaUSE corpus, highlighting similarities and differences between a) assignments and presentations (registers), b) assignments and presentations with respect to each of the four disciplines, c) different sub-registers of assignments and presentations, and d) sub-registers of assignments and presentations with respect to each of the four disciplines. For some features, we will present numerical information, while others will be discussed through non-numerical description.

3.1. Situational characteristics of assignments and presentations

This section explains the similarities and differences in the situational characteristics of assignments and presentations – the two major registers produced by students in Pakistani universities. This comparison is important since one of the aims of the present study is to discover variation in the language produced by Pakistani university students across spoken and written modes as represented by presentations and assignments, respectively. As assignments and presentations are general registers, it is expected that there will be more variation in the situational characteristics of these two registers as compared to specialized registers or sub-registers. The results of the situational analysis of assignments and presentations discussed here are not based on the similarities and differences in the situational features of these two registers in general but are based on the analysis of assignments and presentations included in the PaUSE corpus. The situational characteristics of the assignments and presentations produced by students in Pakistani universities can be divided into three groups of features: a) the features shared by assignments and presentations, a) those that vary across assignments and presentations, and c) those that cannot be specified for general registers like assignments and presentations. Tables 3.1 and 3.2 summarize the situational characteristics of presentations and assignments included in the PaUSE corpus, respectively.

In fact, there are very few situational features that are shared by assignments and presentations represented in the PaUSE corpus. For instance, both assignments and presentations represented in the PaUSE corpus are produced by graduate students, are graded, and the grades count in the overall academic progress. Moreover, all the assignments and presentation included in the PaUSE corpus are judged by the concerned class instructor as fulfilling the minimum passing criteria. Furthermore, there is a possibility of interaction between addressor(s) and audience in both assignments and presentations, however, both differ in their nature of interactiveness. Both assignments and presentations require some specialized knowledge shared between addressor(s) and audience. They address topics related to present time or near past. Another similarity between assignments and presentations lies in their general communicative purpose which, in both cases, is to share information and/or explain concepts and methods. Finally, in both cases, there are far more texts produced by single speaker or writer than groups of speakers or writers.

Overall, the PaUSE corpus contains more assignments than presentations. Out of total texts included in the PaUSE corpus, there are around 63% assignments and 37% presentations. Although assignments and presentations are similar in the sense that most of the texts in each register are produced by single addressors, the two registers differ with respect to the percentage of texts produced by single addressor or groups of addressors. For example, 85% of the total presentations are produced by single speakers or presenters whereas the rest have more than one presenter. On the other hand, a little less than 87% of the assignments are produced by single writer while around 13% have more than one writer. With respect to the number of speaker or writer, assignments and presentations also vary with respect to text length. For instance, presentations made by groups are much longer (1899.31 words per text) than the ones delivered by single speakers (690.33 words per text). On the contrary, assignments produced by single writers are slightly longer (2310.83 words per text) as compared to those written by more than one writer (2254.04 words per text). These differences, of course, have to do with the types of assignments and presentations as well as disciplines which will be discussed in later sections.

Table 3.1. The Situational Characteristics of Presentations in the PaUSE Corpus

Situational Characteristics	Presentations (195 texts)							
	General Presentations (137)				Research-based Presentations (58)			
	BOT (25)	EDU (50)	LING (33)	MS (29)	BOT (3)	EDU (8)	LING (40)	MS (7)
Participants								
Number of speakers								
Single	24	50	21	29	3	7	25	7
Group	1	0	12	0	0	1	15	0
Gender of speakers								
Male	3	26	15	12	0	3	22	3
Female	22	24	16	17	3	5	17	4
Both	0	0	2	0	0	0	1	0
Addressee	Plural							
On-lookers	Yes							
Relations among participants								
Interactiveness	Yes							
Addressor-addressee relationship	Student-teacher / Peers							
Status or power	Low-high / equal							
Shared knowledge	Yes							
Channel								
Mode	Spoken							
Specific medium	Face-to-face, audio-recorded, and transcribed							
Production and comprehension circumstances								
Production	Real-time, planned, cannot be revised, and edited							
Comprehension	Real-time comprehension							
Text length								
Mean text length	674.24	643	1088.09	694.59	415.67	1510.5	1114.98	955.57
Range	141-2054	115-2094	133-2984	116-2487	334-646	981-2295	130-4998	303-2634
Setting								
Place and time	Shared							
Public or private place	Public							
Time	Contemporary							
Communicative purpose								
General Purpose	Present potentially new information / Explain concepts and methods							
Specific purposes	Summarize, incorporate, and synthesize information from multiple sources				Report findings of an original research			
Subject or topic	Various							
Evidence								
Level	Secondary				Primary			
Nature	Qualitative							

Table 3.2. The Situational Characteristics of Assignments in the PaUSE Corpus

Situational Characteristics	Assignments (329 texts)															
	Article Reviews (34)				Literature Reviews (116)				Research Proposals (90)				Research Reports (89)			
	BOT (0)	EDU (6)	LING (21)	MS (7)	BOT (30)	EDU (26)	LING (20)	MS (40)	BOT (14)	EDU (6)	LING (35)	MS (35)	BOT (7)	EDU (25)	LING (36)	MS (21)
Participants																
Number of writers																
Single	0	6	21	7	29	26	9	40	14	6	28	34	5	25	16	19
Group	0	0	0	0	1	0	11	0	0	0	7	1	2	0	20	2
Gender of writers																
Male	0	1	13	5	7	19	12	23	9	3	22	17	5	18	17	14
Female	0	5	8	2	23	7	7	17	5	3	12	18	1	7	18	5
Both	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	2
Addressee	Single															
On-lookers	No															
Relations among participants																
Interactiveness	No															
Addressor-addressee relationship	Student-teacher															
Status or power	Low-high															
Shared knowledge	Yes															
Channel																
Mode	Written															
Specific medium	Typed-written															
Production and comprehension circumstances																
Production	Planned, revised, and edited															
Comprehension	Careful reading or skimming															
Text Length																
Mean text length	0	1138.83	671.71	2362	2720.97	2709.65	1823.75	2356.5	1286	1594.33	1916.06	2404.43	2611.29	2680.64	3348.92	2515.86
Range	0	430-1714	155-1709	1204-3319	252-5927	710-4371	547-3812	356-5472	460-4175	628-2449	307-4512	376-3908	619-6648	819-6641	1076-5816	495-3973
Setting																
Place and time	Not shared															
Public or private place	Private															
Contemporary or historical time	Contemporary															
Communicative purpose																
General Purpose	Present potentially new information / Explain concepts and methods															

Specific purposes	Present a positive and/or negative assessment of a single research article	Summarize, incorporate, and synthesize information from multiple sources	Explain the procedures that are to be followed in a future research	Report findings of an original research			
Subject or Topic			Various				
Evidence							
Level	Secondary	Secondary	Secondary	Primary			
Nature							
Quantitative				3	0	7	15
Qualitative				4	25	29	6

Assignments and presentations also differ regarding the gender of the addressor. For instance, 53.85% of the total presentations are given by female students, 44.61% by male students, and a mere 1.54% involve both genders. On the other hand, most of the assignments are produced by male students (56.23%), followed by those written by female students (41.95%), while only 1.82% of the total assignments are produced by groups involving both males and females. Moreover, presentations produced by groups involving both genders are much longer (1866 words per text) than those produced by either male (865.85 words per text) or female (845.22 words per text) students, while assignments produced by female students (2566.59 words per text) are longer than the ones written by groups involving both genders (2178.5 words per text) and those produced by male students (2110.83 words per text). The final difference between assignments and presentations with respect to participants concerns with addressees and on-lookers. Assignments are produced for single addressee – class instructors or teachers whereas presentations are delivered to plural audience that includes class instructors as well as fellow students or classmates. In fact, it could be argued that presentations, like assignments, are produced to the instructors whereas classmates listening to the presentations are more like on-lookers. The rationale behind this argument is that the purpose of giving presentations is to convince or persuade the instructor that one has grip or command on a topic based on library or field research and one is capable of delivering the content related to a topic in front of the audience in spoken mode so as to get high grades from the instructors, that is why, the instructor is the real addressee of the presentations. Although a secondary goal of presentations is to educate or inform fellow students on a certain topic; however, since they do not have the power to judge the presentations as pass or fail, they are addressed only indirectly by the presenters.

Assignments and presentations included in the PaUSE corpus also vary with respect to relations among participants. First, presentation, although consist of monologic speech delivered to a group of audience, has the potential to be interactive as instructor as well as audience can ask questions or request for clarification during or after the presentations. However, while transcribing the presentations included in the PaUSE corpus, it was revealed that there was very little interaction between addressor(s) and audience, and if there were any questions, they were a) initiated by the instructor and b) asked and responded to in students' first language. Students' not asking questions was seen as a cooperative gesture as it is usually thought that if the presenter is unable to provide proper response to a question, it will have negative effect on his or her grades.

Assignments, unlike many written registers, have some element of interactivity as instructors can provide feedback on first and subsequent drafts of an assignment. Second, there is student-teacher relationship between addressor and addressee in case of assignments while there is student-teacher as well as student-student relationship between the participants in the context of presentations. Finally, addressee has more power or status than addressor in assignments whereas in presentations, all the addressees except the instructor have a status equal to that of the addressor.

The most obvious and telling difference between assignments and presentations lies in channel of communication and in production and comprehension circumstances. Assignments included in the PaUSE corpus are written texts produced either in typed form and sent to us via e-form or in printed form which were later scanned and converted into .txt files. These, like most other written registers, are usually planned, revised, and edited. However, in the context of Pakistani universities, students usually do not submit multiple drafts of an assignment; rather, they submit only final draft without receiving any feedback from the teachers as they are habitual of completing the first draft of assignments just before the assignment is due. In this case, Pakistani university students' assignments resemble open-book exam with ample time for planning and writing. The instructors might read the assignments carefully or skim through them focusing on important parts. On the other hand, presentations represented in the PaUSE corpus were produced in spoken form and were audio-recorded and later transcribed; however, these differ from most interpersonal spoken registers with respect to their production circumstances. Unlike typical interpersonal spoken registers such as face-to-face conversation which is produced in real time and is often not planned, presentations included in the PaUSE corpus are planned and often accompany a written text which could either be the actual written assignment or handouts or slides. Many presentations represented in the PaUSE corpus involve reading from the slides or handouts rather than presenting the content in speakers' own language. Reading from the slides or handouts is not unusual in case of presentations made by non-native speakers of English as they are intimidated because of lack of any training related to giving presentations as reported by Hyland (2009) and Zareva (2009a). In this regard, student presentations bear a close association with classroom lectures as both are planned and accompany some written texts. Like lectures (Csomay, 2006), student presentations can also be considered at an interface of oral-literate continuum with presentations bearing

resemblance to both face-to-face conversation and academic writing. As presentations are usually based on written texts such as library research, empirical research, lab report, or simply handouts or slides; it is expected that some of the linguistic features employed by the presenters would be similar to those typically found in academic prose. On the other hand, presentations, like face-to-face conversation, are produced in real time, cannot usually be revised or edited during their delivery, and require read time comprehension on the part of the addressee(s), they show some resemblance with oral discourse.

Assignments and presentations included in the PaUSE corpus also vary in terms of text length and setting. On average, the assignments, as expected, are much longer (2303.23 words per text) than the presentations (870.13 words per text). The longest assignment contains 6648 words while the longest presentation contains 4998 words. On the other hand, the shortest assignment consists of only 155 words while the shortest presentation has only 115 words. Thus, the range (maximum words minus minimum words) of assignments (6493) is greater than that of presentations (4883). The number of words in shortest assignments and presentations is surprising which will be explained while discussing the situational characteristics of assignments and presentations with respect to sub-registers and disciplines.

Finally, in case of presentations, addressors and addressees share time and place (classroom) as the interaction between them is face-to-face. On the contrary, the addressors and addressees of the assignments do not share the same time or space, although it is expected that the instructors would read the assignments a few days after they were produced. The time, in this sense, can be considered somewhat shared. In this regard, assignments bear some resemblance with newspaper writing which is expected to be read on the same day or a day after it is produced. The place of communication of assignments is private as the communication only happens between student and class instructor.

Assignments and presentations cannot be compared on certain situational features such as specific communicative purpose(s), subject or topic, and level and type of evidence since there is variation within assignments and presentations on these features with respect to sub-registers as well as disciplines.

3.2. Situational characteristics of assignments and presentations with respect to disciplines

This section describes variation in the situational characteristics of assignments and presentations with respect to four disciplines. Compared to the large number of differences in the situational characteristics of assignments and presentations included in the PaUSE corpus, there are only four situational features on which assignments and presentations vary with respect to disciplines: number of addressor(s), gender of addressor(s), text length, and general topic or subject.

With respect to number of addressor(s), almost all the assignments and presentations from three out of four disciplines, except linguistics, are produced by single speaker or writer. Although linguistics also follows the same trend, there are considerable number of assignments and presentations from linguistics that are produced by more than one speaker or writer. Around 37% of the presentations and around 34% of the assignments from linguistics are produced by groups of addressors.

As far as the gender of the addressor(s) is concerned, in two out of four disciplines – botany and linguistics – members of the same gender group produced more assignments as well as presentations. For instance, in case of botany, female students produced more assignments (29 texts) and presentations (22 texts) than those produced by male students (21 assignments and 6 presentations). From the discipline of linguistics, on the other hand, males produced more assignments (64 texts) as well as presentations (37 texts) as compared to the ones produced by females (45 assignments and 33 presentations). Opposite trend is found in management sciences where one gender group produced more assignments while the other gave more presentations. The PaUSE corpus contains more presentations from management sciences that are given by female students (21 texts) than those given by male students (15 texts), whereas it contains more assignments from management sciences that are produced by male students (59 texts) than those of female students (42 texts). Finally, in case of education, male students produced almost twice as many assignments (41 texts) as were produced by female students (22 texts) while the PaUSE corpus contains an equal number of presentations from education written by the members of both gender groups (29 texts each).

Assignments and presentations in all four disciplines also vary with respect to mean text length measured in terms of average number of words per text, and range which is measured as the difference between maximum

and minimum values. One common trend across disciplines is that assignments in all four disciplines are longer and have broader range than presentations. However, if we consider variation within assignments in terms of mean text length, the assignments from education (2442.32 words per text) are longer than those produced by the students of management sciences (2405.65 words per text), botany (2312 words per text), and linguistics (2126.82 words per text). Moreover, assignments from botany have the broadest range (6396 words), followed by those from education (6211 words), linguistics (5661 words), and management sciences (5116 words), respectively. This means that all four disciplines represented in the PaUSE corpus contain very long texts as well as very short ones. On the other hand, students from linguistics (1102.82 words per text) delivered longer presentations than those given by the students of education (762.66 words per text), management sciences (745.33 words per text), and botany (646.54 words per text). Furthermore, presentations from linguistics display the broadest range (4868 words), followed by the presentations from management sciences (2518 words), education (2180 words), and botany (1913 words), respectively.

Finally, there is variation in assignments and presentations with respect to disciplines in their subject or topic. Both assignments and presentations from botany included in the PaUSE corpus focus on plants, natural resources, environment, etc. The focus of assignments and presentations from education, on the other hand, is on theoretical, cognitive, and methodological issues related to teaching, assessment, curriculum development, etc. at various levels. The focus of assignments and presentations from linguistics is on studying English and local languages from various vantage points, that is, phonological, morphological, syntactic, second language acquisition, pragmatics, etc. Finally, the spoken and written texts from management sciences are geared towards various issues related to marketing, finance, human resource management, entrepreneurship, business communication, etc. These are general topic areas from each discipline. It is not possible to describe specific topics of each assignment and presentation since the PaUSE corpus contains as many topics as there are texts.

3.3. Situational characteristics of text-types or sub-registers

This section describes variation in the situational characteristics of sub-registers or text-types represented in the PaUSE corpus. The PaUSE corpus contains six sub-registers which include four sub-registers of assignments and two of presentations. The four sub-registers of assignments

are article reviews, literature reviews, research proposals, and research reports, whereas the two sub-registers of presentations are general presentations and research-based presentations. These six sub-registers vary on the following situational features: number of addressor(s), gender of addressor(s), mean text length in terms of average number of words per text and range, specific communicative purpose(s), and level of evidence.

Out of six sub-registers represented in the PaUSE corpus, article review is the only text type that is produced by individual or single writers; in all other types, there are texts produced by single as well as groups of addressors. In each of the remaining five sub-registers, single addressor texts are more in number than those produced by more than one speaker or writer. However, the sub-registers differ with respect to the percentage of texts produced by more than one addressor. For instance, irrespective of spoken and written modes, the two types of sub-registers involving actual research – research-based presentations and research reports – have greater percentage of texts produced by groups of addressors as compared to the sub-registers involving library research, that is, 27.59% of the total research-based presentations and 26.97% of the total research reports are produced by more than one writer or speaker, whereas in case of literature reviews, general presentations, and research proposals, the texts produced by more than one addressor occur in 10.34%, 9.49%, and 8.89%, respectively. If we consider the gender of the addressor(s) in connection with sub-registers, general presentations are the only type of sub-register in which more texts are produced by female students (57.66%) as compared to male students (40.88%). On the contrary, although male students produced more texts in each of the remaining five sub-registers, there is variation in the percentage of each of the sub-registers produced by male and female students. For instance, in four out of five sub-registers – research-based presentations (44.83 vs 53.45%), article reviews (44.12% vs 55.88%), literature reviews (46.55% vs 52.59%), and research proposals (42.22% vs 56.67%) – the ratio between the texts produced by male and female students is almost the same, whereas the percentage of research reports produced by female students is much less (34.83%) than those produced by male students (60.67%).

Text length is yet another situational feature on which sub-registers included in the PaUSE corpus vary. Irrespective of spoken and written modes, research reports in the PaUSE corpus are longer than the other five sub-registers with respect to the average number of words per text. These occur with the mean length of 2906.62 words per text, followed by literature

reviews which have the mean length of 2438.06 words per text, research proposals with the mean length of 1986.52 words per text. Surprisingly, research-based presentations, on average, are slightly longer (1114.12 words per text) than article reviews (1102.15 words per text), whereas general presentations with the mean length of 766.83 are the shortest amongst all sub-registers represented in the PaUSE corpus. The most surprising aspect concerning the lengths of sub-registers is that some of the texts belonging to each sub-register are unexpectedly very short. For instance, the shortest text in the PaUSE corpus containing mere 115 words belongs to the sub-register of general presentations, while the shortest texts coming from the categories of research-based presentations, article reviews, literature reviews, research proposals, and research reports contain 130, 155, 252, 307, and 495 words, respectively. Some texts, on the other hand, from each sub-register are exceptionally longer than mean text length for each category. For instance, the longest texts belonging to the sub-registers of research reports, literature reviews, research-based presentations, research proposals, article reviews, and general presentations contain 6648, 5927, 4998, 4512, 3319, and 2984 words, respectively. The nature of very short and very long text from each sub-register will be explained with reference to discipline in the next section.

The sub-registers or text types also differ with respect to their communicative purposes. Research reports and research-based presentations, despite having different modes of production, have similar communicative purpose, that is, both report findings of an original research related to a field of study or explain the rationale, methods, and results of an original research. Moreover, general presentations and literature reviews, which are both based on library research, also have somewhat similar communicative purpose which is to investigate an issue by summarizing, incorporating, and synthesizing information from multiple sources. Furthermore, research proposals are somewhat related to research reports in the sense that the focus of both sub-registers is on original research, however, contrary to the communicative purpose of research reports - presenting findings of a research already conducted -, research proposals explain the procedures that are to be followed in a future study. On the other hand, the communicative purpose of article reviews is different from that of all other sub-registers, that is, to present a positive and/or negative assessment of a single research article. Contrary to the involvement of multiple sources in case of other five text-types irrespective of their mode of production, the focus of article reviews is on a single article. Finally, all six sub-registers included in the PaUSE corpus, irrespective of spoken and

written mode, share one common characteristic: they all contain reference to previous literature. However, the sub-registers differ from each other with respect to the extent to which previous literature is cited, with article review containing only one study and others incorporating multiple sources. The sub-registers also vary internally with reference to the amount of previous literature included. Furthermore, research reports, research-based presentations, and research proposals are similar in the sense that all contain description of the methods used or will be used to collect and analyze data.

Finally, the sub-registers included in the PaUSE corpus also differ in terms of level of evidence, that is, whether they contain primary evidence or secondary evidence. Research reports and research-based presentations are similar in their use of primary evidence, whereas general presentations, literature reviews, and article reviews contain secondary sources with the difference that article reviews only contain single or a handful of secondary sources while the other two usually are based on evidence from multiple sources. Finally, research proposals are different from all other sub-registers in the sense that although these do not contain primary evidence, however, since the purpose of research proposals is to present roadmap for future research including how data will be collected and analyzed, in this way, it can be said that it contains primary evidence.

3.4. Situational characteristics of sub-registers with respect to disciplines

In this section, we will present the differences in the situational characteristics of each sub-register in connection with four disciplines represented in the PaUSE corpus. The characteristics of each sub-register in all four disciplines will be described on the following situational features: number of speaker or writer, gender of speaker or writer, text length.

General presentations from three out of four disciplines, excluding linguistics, are produced by single speakers with only one presentation from botany is given by more than one speaker. From linguistics, 36.36% of the total general presentations are made by more than one speaker. Furthermore, in three out of four disciplines, except education, more general presentations in the PaUSE corpus are produced by female students with botany having the highest percentage of general presentations produced by female students (i.e., 88%). However, the percentage of general presentations made by male students from education (52%) is slightly

higher than those produced by female students (48%). Moreover, general presentations made by the students of linguistics (1088.09 words per text) are much longer than those produced by the students of management sciences (694.59), botany (674.24 words per text), and education (643 words per text). Although all four disciplines have some fairly long general presentations presenting exhaustive review of literature on a given topic or issue, as indicated by maximum number of words in Table 3.1, it is very surprising that the shortest texts from each of the four disciplines are very short, that is, less than 150 words. The content of these presentations is mainly responsible for their short length. For instance, the shortest general presentations from botany and linguistics consist only of different definitions of a type of plant (botany) and two types of linguistic phenomenon, that is, phrasal movement and head movement (linguistics). On the other hand, the shortest general presentations from management sciences contain some examples related to gestures and postures as a type of non-verbal communication. Finally, the two shortest general presentation from education explain two other reasons for the short length of general presentations: code-switching and shortage of time. In case of both presentations, most of the content is delivered in Urdu and only the content presented in English is transcribed. Moreover, one of the presenters is not allowed to complete her presentation because of shortage of time as the class came to an end.

The second sub-register included in the PaUSE corpus – research-based presentations – contains most of the texts from linguistics with only a handful of texts from the other three disciplines. Research-based presentations include presenting the report of a completed research paper (based on research reports) as well as describing the procedures that will be followed for conducting a future study (based on research proposals). Although most of the research-based presentations are given by single speakers, 37.5% of the research-based presentations from linguistics are produced by more than one speaker. As far as the gender of addressors is concerned, research-based presentations from education and management sciences produced by female students (62.5% and 57.14% respectively) are slightly more in number as produced by male students, while the research-based presentations coming from botany are exclusively produced by female students. However, in case of linguistics, the PaUSE corpus includes more texts produced by male students (55%) as compared to female students (42.5%). In terms of average number of words per text, research-presentations from education (1510.5 words) are the longest, followed by those from linguistics (1114.98 words), management sciences (955.57

words), and botany (415.67 words). Some of the research-based presentations are very long whereas some are very short. The length of these presentations is directly proportional to the relative importance given to some part(s) of the research. For instance, the longer research-based presentation present detailed accounts of research projects or reports describing introduction, literature review, methods, and/or results. Whereas the focus of presenters in case of shorter presentations is on some specific part. Apart from education, the other three disciplines include some short research-based presentations although only one of them contains less than 150 words. The shortest research-based presentations from botany and management sciences only present the results, while the one containing minimum number of words from linguistics is an incomplete presentation since the class instructor interrupted the presenters by asking questions right from the beginning and never let them complete their presentation.

Turning to the written sub-registers, article reviews mainly come from linguistics (61.76%) with a few texts from education (17.65%) and management sciences (20.59%) too while the PaUSE corpus does not contain any article review from botany. All the article reviews included in the PaUSE corpus were produced by single writers. Female students produced more article reviews from education (83%) whereas article reviews produced by male students from linguistics (61.90%) and management sciences (71.43%) are greater in number than those written by female students. On average, article reviews produced by the students of management sciences (2362 words per text) are longer than those produced by the students of education (1138.83 words per text) and linguistics (671.71). The longer article reviews include summaries of the articles as well as detailed comments whereas the shorter ones only contain comments on an article. For instance, the shortest article review is from linguistics which does not include summary of the article but contains only a few comments on the article.

Literature reviews in three out of four disciplines, excluding linguistics, are mainly produced by single authors. Surprisingly, there are more literature reviews from linguistics produced by multiple authors (55%) as compared to single-author literature reviews (45%). Most of the literature reviews from botany are produced by female students while male students produced more literature reviews as compared to female students from the other three disciplines. On average, literature reviews coming from botany are the longest (2720.97 words per text), closely followed by those from education (2709.65 words per text), and then management sciences

(2356.5 words per text), and linguistics (1823.75 words per text). Some of the literature reviews from each discipline are very long incorporating and synthesizing a wide range of literature related to a topic or issue while the shorter ones contain a few sources and/or just summarize the findings of some sources. Moreover, many literature reviews do not include in-text citations; they only have a reference list at the end and thus resemble more with essays, whereas others contain in-text citations and no reference lists, while still others contain both reference lists as well as in-text citations.

As far as the sub-register of research proposal is concerned, in three out of four disciplines, barring linguistics, almost all the research proposals included in the PaUSE corpus are produced by single writers with one research proposal from management sciences coming from more than one writer. Similarly, although majority of the research proposals from linguistics also come from single authors, 20% of the total research proposals from linguistics are produced by more than one writer. Furthermore, male students produced more research proposals from botany and linguistics than female students whereas male and female students produced equal number of research proposals from education, and female students produced slightly more research proposals than male students in case of management sciences. In terms of average number of words per text, the longest research proposals included in the PaUSE corpus come from management sciences (2404.43 words per text), followed by those from linguistics (1916.10 words per text), education (1594.33 words per text), and botany (1286 words per text). As with other sub-registers, all four disciplines contain some very long research proposals. These long research proposals, irrespective of discipline, have detailed literature review including objectives and research questions, methodology, and proposed analytical procedures. Contrary to this homogeneity, the shorter research proposals from each discipline differ with respect to the parts focused. For instance, the shorter research proposals from botany only contain short introduction with no literature review, objectives, and very brief methodology, whereas those from education contain brief introduction, objectives, research questions, hypothesis, and a very brief description of framework. The shorter research proposals from linguistics are more like brief synopsis with short mention of objectives, research questions, and methods, while those from management sciences have very brief introduction followed by short description of methods.

Finally, in case of the sub-register of research reports represented in the PaUSE corpus, education is the only discipline with all single-author

research reports while the other three disciplines contain research reports produced by both single and groups of writers. However, the three disciplines differ with respect to the percentage of research reports produced by single versus group authors. The PaUSE corpus contains a lot more single-author research reports from management sciences and a slightly more single-authored research reports from botany whereas more texts included in the PaUSE corpus from linguistics come from multiple writers (55.55%) as compared to those that come from single writers (44.44%). As far as the gender of the writer is concerned, female students from linguistics produced slightly more research reports than those produced by male students whereas in case of other three disciplines, research reports produced by male students are more in number than those written by female students. Moreover, in terms of average number of words per text, research reports from linguistics (3348.92 words per text) are the longest, followed by almost similar mean text length for those from education (2680.64 words per text), botany (2611.29 words per text), and management sciences (2515.86 words per text). Furthermore, the longer texts from all four discipline have one common characteristic: their format closely resembles the format of a published research article, and they contain all important parts of a published research paper such as introduction, methods, results, discussion, etc. The shorter texts from each discipline do not contain all these parts but mainly focus on methods and/or results. For instance, although the shorter research reports from botany, linguistics, and education contain most parts of a published research article, the explanation of each part is very brief and these shorter research reports are replete with tables and figures which were removed during the compilation of the PaUSE corpus which resulted in reducing the text length. The shorter texts from management sciences, on the other hand, only consist of very brief methods section followed by very short description of results. Finally, although all research reports from all four disciplines made use of primary evidence or data, the disciplines differ with respect to the use of quantitative versus qualitative evidence. Research papers included in the PaUSE corpus predominantly use qualitative analysis (around 80%), those from three disciplines excluding education also contain reports (around 20%) presenting results in quantitative terms. However, the percentage of quantitative reports in each of these disciplines is different with most research reports from management sciences making use of quantitative evidence (71.43%), followed by those from botany (42.86%) and linguistics (19.44%). These quantitative reports, besides presenting results of descriptive statistical measures, use variety of inferential statistical tests including t-test, Analysis of Variance (ANOVA), correlation, factor

analysis, etc. Moreover, some of the research reports that use qualitative evidence also present result in the form of percentages and frequencies, however, these are not treated as having quantitative evidence.

4. Conclusion

This study sought to explore variation in Pakistani university student English on situational characteristics with respect to four variables: registers, registers across disciplines, sub-registers, and sub-registers across disciplines. The findings clearly show that although there is variation in the situational characteristics of Pakistani university student English with respect to all four variables, there is more variation in the situational characteristics of the two main registers (presentations and assignments) due to different modes of production (spoken versus written) than in registers across disciplines, sub-registers, and sub-registers across disciplines. The two main registers differ with respect to the number and length of texts in each register, characteristics of addressor(s) with respect to number and gender, characteristics of addressees, relationship among addressor(s) and addressees, channel, production and comprehension circumstances, and time and place of communication. Moreover, presentations and assignments across disciplines differ with respect to the number and gender of addressor(s), text length, and topic or subject. Furthermore, the sub-registers included in the PaUSE corpus vary with respect to the number and gender of addressor(s), text length, communicative purpose, and level of evidence. Finally, the sub-registers also vary across disciplines with respect to text length, and number and gender of speakers or writers. The variation in the situational characteristics of Pakistani university student English makes a strong case for exploring variation in the linguistic characteristics of academic English produced by students in Pakistani universities which will be presented in future studies.

References

Abdulaziz, M. (2017). *A multi-dimensional analysis of Pakistani learner writing* (Unpublished doctoral dissertation). Government College University, Faisalabad.

Abdulaziz, M., Mahmood, M. A., &Azher, M. (2016). Variation in learner's argumentative essays: A multidimensional comparative analysis. *Science International, 28*(4), 413-424.

Adel, A., &Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes, 31*, 81-92.

Afsar, A., & Kamran, U. (2011). Comparing consonantal phonemes of Pakistani Standard English with British Standard English. *Kashmir Journal of Language Research, 14*(1), 29-48.

Ahmad, S., & Mahmood, M. A. (2015). Comparing explicit features of Pakistani press reportage with British press reportage: A multi-dimensional analysis. *NUML Journal of Critical Inquiry, 13*(2), 1-31.

Ahmad, S., & Mahmood, M. A. (2016). Linguistic variation among newspapers in Pakistani print Media: A multidimensional analysis. *Kashmir Journal of Language Research,19*(1), 135-156.

Ali, S., & Ahmad, S. (2016). Discourse style variation among the leading novelists of Pakistani fiction in English: A multidimensional analysis. *ELF Annual Research Journal, 18*, 227-246.

Ali, S., & Ahmad, S. (2017). Discourse style of Pakistani fiction in English: A multidimensional analysis. *Balochistan Journal of Linguistics, 5*, 9-28.

Ali, S., & Ahmad, S. (2019). Bilingual creativity and world Englishes: A comparative multidimensional analysis of Pakistani English general fiction with British general fiction. *Kashmir Journal of Language Research, 22*(1), 121-144.

Ali, S., & Shakir, M. A. (2016). Style variation among the sub-genres of Pakistani fiction in English: A multidimensional analysis. Journal *of Social Sciences, 7*(2), 89-114.

Alvi, U., Mehmood, M. A., & Rasool, S. (2016). A multidimensional analysis of Pakistani press editorials. *The Dialogue, 11*(3), 270-284.

Asghar, S. A., Mahmood, M. A., & Asghar, Z. (2018a). A multidimensional analysis of Pakistani legal English. *International Journal of English Linguistics, 8*(5), 215-229.

Asghar, S. A., Mahmood, M. A., & Asghar, Z. (2018b). Linguistic variation across Pakistani legal genres: A multidimensional analysis. *ELF Annual Research Journal, 20*, 133-158.

Barbieri, F. (2015). Involvement in university classroom discourse: Register variation and interactivity. *Applied Linguistics, 36*(2), 151-173.

Baumgardner, R. (1990). The indigenization of English in Pakistan. *English Today*, *6*(1), 59-65.

Baumgardner, R. (1993). Utilizing Pakistani newspaper English to teach grammar. In R. Baumgardner (Ed.), *The English language in Pakistan* (pp. 255-273). Karachi: Oxford University Press.

Baumgardner, R. (1995). Pakistani English: Acceptability and the norm. *World Englishes, 14*(2), 261-71.

Baumgardner, R. (1998). Word-formation in Pakistani English. *English World-Wide, 19*(2), 205-246.

Baumgardner, R., Kennedy, A., & Shamim, F. (1993). The Urduization of English in Pakistan. In R. Baumgardner (Ed.), *The English language in Pakistan* (pp. 83-203). Karachi: Oxford University Press.

Begum, M. (2016). *Linguistic variation across academic writing of Pakistani university students: A multidimensional analysis* (Unpublished doctoral dissertation). Government College University, Faisalabad.

Biber, D. (1988). *Variation across speech and writing.* Cambridge: Cambridge University Press.

Biber, D. (1994). An analytical framework for register studies. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 31-56). New York: Oxford University Press.

Biber, D. (2006a). Stance in spoken and written university registers. *Journal of English for Academic Purposes, 5*, 97-116.

Biber, D. (2006b). *University language: A corpus-based study of spoken and written registers.* Amsterdam: John Benjamins.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes, 26*, 263-286.

Biber, D., & Conrad, S. (2009). *Register, genre, and style.* New York: Cambridge University Press.

Biber, D., Conrad, S., & Cortes, V. (2004). *If you look at…*: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371-405.

Biber, D., Conrad, S., Reppen, R., Byrd, P., &Helt, M. (2002). Speaking and writing in the university: A multidimensional comparison. *TESOL Quarterly, 36*(1), 9-48.

Biber, D., Conrad, S., Reppen, R., Byrd, P., Helt, M., Clark, V., … Urzua, A. (2004). *Representing language use in the university: Analysis of the TOEFL 2000 spoken and written academic language corpus* (TOEFL Monograph Series No. MS-25). Princeton, NJ: Educational Testing Service.

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing.* Cambridge: Cambridge University Press.

Cao, Y., & Xiao, R. (2013). A multi-dimensional contrastive study of English abstracts by native and non-native writers. *Corpora, 8*(2), 209-234.

Carkin, S. (2001). *Pedagogic discourse in introductory classes: Multi-dimensional analysis of textbooks and lectures in biology and*

macroeconomics (Unpublished doctoral dissertation). Northern Arizona University, Flagstaff.

Chen, Y.-H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology, 14*(2), 30-49.

Conrad, S. (1996a). *Academic discourse in two disciplines: Professional writing and student development in biology and history* (Unpublished doctoral dissertation). Northern Arizona University, Flagstaff.

Conrad, S. (1996b). Investigating academic texts with corpus-based techniques: An example from biology. *Linguistics and Education, 8*, 299-326.

Conrad, S. (2018). The use of passives and impersonal style in civil engineering writing. *Journal of Business and Technical Communication, 32*(1), 38-76.

Cooper, A., & Bikowski, D. (2007). Writing at the graduate level: What tasks do professors actually require? *Journal of English for Academic Purposes, 6,* 206-221.

Cortes, V. (2002). Lexical bundles in freshman composition. In R. Reppen, S. Fitzmaurice, & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 131-145). Amsterdam: John Benjamins.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23*, 397-423.

Csomay, E. (2002). Variation in academic lectures: Interactivity and level of instruction. In R. Reppen, S. Fitzmaurice, & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 203-224). Amsterdam: John Benjamins.

Csomay, E. (2005). Linguistic variation within university classroom talk: A corpus-based perspective. *Linguistics and Education, 15*, 243-274.

Csomay, E. (2006). Academic talk in American university classrooms: Crossing the boundaries of oral-literate discourse. *Journal of English for Academic Purposes, 5*, 117-135.

Csomay, E. (2007). A corpus-based look at linguistic variation in classroom interaction: Teacher talk versus student talk in American university classes. *Journal of English for Academic Purposes, 6*, 336-355.

Csomay, E. (2015). A corpus-based analysis of linguistic variation in teacher and student presentations in university settings. In V. Cortes & E. Csomay (Eds.), *Corpus-based research in applied linguistics: Studies in honor of Doug Biber* (1-24). Philadelphia: John Benjamins Publishing Company.

Edwards, J. A., & Lampert, M. D. (Eds.). (1993). *Talking data: Transcription and coding indiscourse research*. Hillsdale: Lawrence Erlbaum.

Egbert, J. (2014). *Reader perceptions of linguistic variation in published academic writing* (Unpublished doctoral dissertation). Northern Arizona University, Flagstaff.

Fortanet, I. (2004). The use of 'we' in university lectures: Reference and function. *English for Specific Purposes, 23*, 45-66.

Friginal, E., & Mustafa, S. S. (2017). A comparison of U.S.-based and Iraqi English research article abstracts using corpora. *Journal of English for Academic Purposes, 25*, 45-57.

Gardner, S., Nesi, H., & Biber, D. (2019). Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing. *Applied Linguistics, 40*(4), 646-674.

Gray, B. (2010). On the use of demonstrative pronouns and determiners as cohesive devices:

A focus on sentence-initial *this/these* in academic prose. *Journal of English for Academic Purposes, 9*, 167-183.

Gray, B. (2011). *Exploring academic writing through corpus linguistics: When discipline tells only part of the story* (Unpublished doctoral dissertation). Northern Arizona University, Flagstaff.

Gray, B., & Cortes, V. (2011). Perception vs. evidence: An analysis of *this* and *these* in academic prose. *English for Specific Purposes, 30*, 31-43.

Hardy, J. A. (2015). Multi-dimensional analysis of academic discourse. In P. Baker & T. McEnery (Eds.), *Corpora and discourse studies: Integrating discourse and corpora* (pp. 155-174). Hampshire: Palgrave Macmillan.

Hardy, J. A., &Friginal, E. (2016). Genre variation in student writing: A multi-dimensional analysis. *Journal of English for Academic Purposes, 22*, 119-131.

Hardy, J. A., &Römer, U. (2013). Revealing disciplinary variation in student writing: A multi-dimensional analysis of the Michigan corpus of upper-level student papers (MICUSP). *Corpora, 8*(2), 183-207.

Hussain, Z., Mahmood, M. A., &Azher, M. (2016). Register variation in Pakistani English: A multidimensional approach. *Science International, 28*(4), 391-402.

Hyland, K. (2009). *Academic discourse: English in a global context*. London: Continuum. Iberri-Shea, G. (2011). Speaking in front of the class: A multi-dimensional comparison of

university student public speech and university language. *Classroom Discourse, 2*(2), 251-267.

Jabeen, F., Mahmood, M. A., & Rasheed, S. (2011). An attitudinal study of Pakistani English. *Interdisciplinary Journal of Contemporary Research in Business, 3*(5), 109-119.

Jabeen, F., Rai, M. A., &Arif, S. (2011). A corpus based study of discourse markers in British and Pakistani speech. *International Journal of Language Studies, 5*(4), 69-86.

Kachru, B. (1983). *The Indianization of English: The English language in India.* New Delhi: Oxford University Press.

Kachru, B. (1988). The sacred cows of English. *English Today, 16*(1), 3-8.

Kennedy, A. (1993). A bribe by any other name…: Terms of gratification in Pakistani English. In R. Baumgardner (Ed.), *The English language in Pakistan* (pp. 204-211). Karachi: Oxford University Press.

Leedham, M., & Fernandez-Parra, M. (2017). Recounting and reflecting: The use of first person pronouns in Chinese, Greek and British students' assignments in engineering. *Journal of English for Academic Purposes, 26*, 66-77.

Mahboob, A. (2009). English as an Islamic language: A case study of Pakistani English. *World Englishes, 28*(2), 175-189.

Mahboob, A., & Ahmar, N. H. (2004). Pakistani English: Phonology. In E. W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie, & C. Upton (Eds.), *A handbook of varieties of English: A multimedia reference tool*. New York: Mouton de Gruyter.

Mahmood, M. A. (2009). *A corpus based analysis of Pakistani English* (Unpublished doctoral dissertation). Bahauddin Zakariya University, Multan.

Mahmood, M. A., & Hussain, Z. (2016). Linguistic variation across written registers of Pakistani English: A multidimensional analysis. *Pakistan Journal of Languages and Translation Studies, 4*, 15-36.

Mahmood, R. (2009). *A lexico-grammatical study of noun phrase in Pakistani English* (Unpublished doctoral dissertation). Bahauddin Zakariya University, Multan.

Malik, S., & Saeed, T. (2015). Lack of aspiration in unvoiced plosives in Pakistani English: An acoustic analysis based study. *Kashmir Journal of Language Research, 18*(3), 87-103.

Nesi, H., &Basturkmen, H. (2006). Lexical bundles and discourse signaling in academic lectures. *International Journal of Corpus Linguistics, 11*(3), 283-304.

O'Boyle, A. (2014). 'You' and 'I' in university seminars and spoken learner discourse. *Journal of English for Academic Purposes, 16, 40-56.*

Peacock, M. (2011). A comparative study of introductory *it* in research articles across eight disciplines. *International Journal of Corpus Linguistics, 16*(1), 72-100.

Poos, D., & Simpson, R. (2002). Cross-disciplinary comparisons of hedging: Some findings from the Michigan corpus of academic spoken English. In R. Reppen, S. Fitzmaurice, & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 3-23). Amsterdam: John Benjamins.

Qasim, S., & Shakir, A. (2016). Linguistic variation of Pakistani fiction and non-fiction book blurbs: A multidimensional analysis. *ELF Annual Research Journal, 18*, 185-206.

Rahman, T. (2010). *Pakistani English: The linguistic description of a non-native variety of English* (2nd ed.). Islamabad: National Institute of Pakistan Studies, Quaid-i-Azam University.

Rashid, A., & Mahmood, M. A. (2019). Linguistic variations across disciplines: A multidimensional analysis of Pakistani research articles. *Global Social Sciences Review, 4*(1), 23-33.

Rashid, A., Mahmood, M. A., & Ahmad, S. (2017). Linguistic variation across research sections: A multidimensional analysis of Pakistani academic journal articles. *Global Language Review, 2*, 15-37.

Romer, U., & O'Donnell, M. B. (2011). From student hard drive to web corpus (part 1): The design, compilation and genre classification of the Michigan corpus of upper-level student papers (MICUSP). *Corpora, 6*(2), 159-177.

Shakir, A. (2013). *Linguistic variation across print advertisements in Pakistani media: A multidimensional analysis* (Unpublished doctoral dissertation). International Islamic University, Islamabad.

Shakir, M., &Deuber, D. (2018). A multidimensional study of interactive registers in Pakistani and US English. *World Englishes, 37*, 607-623.

Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication, 33*(2), 149-183.

Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing, 32*, 17-35.

Talaat, M. (1993). Lexical variation in Pakistani English. In R. Baumgardner (Ed.), *The English language in Pakistan* (pp. 55-62). Karachi: Oxford University Press.

Tasker, D. G. (2019). *Situational and linguistic variation in undergraduate English-department student writing* (Unpublished doctoral dissertation). Northern Arizona University, Flagstaff.

Weigle, S. C., &Friginal, E. (2015). Linguistic dimensions of impromptu test essays compared with successful student disciplinary writing: Effects of language background, topic, and L2 proficiency. *Journal of English for Academic Purposes, 18*, 25-39.

Yousaf, M., & Shehzad, W. (2018). Fixedness of expressions in doctoral research dissertations: A corpus based analysis. *Kashmir Journal of Language Research, 21*(2), 27-45.

Zareva, A. (2009a). Informational packaging, level of formality, and the use of circumstance adverbials in L1 and L2 student academic presentations. *Journal of English for Academic Purposes, 8*, 55-68.

Zareva, A. (2009b). Student academic presentations: The processing side of interactiveness. *English Text Construction, 2*(2), 265-288.

Zareva, A. (2011). 'And so that was it': Linking adverbials in student academic presentations. *RELC Journal, 42*(1), 5-15.

Zareva, A. (2012). Lexical composition of effective L1 and L2 students' academic presentations. *Journal of Applied Linguistics, 6*(1), 91-110.

Zareva, A. (2016). Multi-word verbs in student academic presentations. *Journal of English for Academic Purposes, 23*, 83-98.