Features Characterizing Academic Writing: A Corpus-Based Research on Dissertations from Hard and Soft Sciences Disciplines

Muhammad Ahmad (Corresponding Author)[1], Muhammad Asim Mahmood[2], Ali Raza Siddique[3]

Keywords:

Academic writing,

Clausal features,

Intermediate features,

Lexicogrammatical features,

Phrasal features.

Abstract

This corpus-based research aims to find out the salient features characterizing Pakistani academic writing (AW). In this regard, the corpus has been: developed from doctoral dissertations from hard and soft sciences disciplines representing Pakistani AW; tagged through MAT and TagAnt Taggers; and analyzed through AntConc Software. Results show the use of clausal, phrasal, and intermediate features more in the AW from soft sciences than in the AW from hard sciences. Therefore, phrasal, intermediate, and clausal features are concluded to characterize AW from soft sciences discipline more than AW from hard sciences discipline. Furthermore, the use of phrasal, intermediate, and clausal features is found to contain in the first, second, and third more frequent use. Therefore, phrasal features are concluded to characterize Pakistani AW. In addition, phrasal features are found to characterize AW from soft sciences discipline more than the AW from hard sciences discipline. Therefore, it is concluded that Pakistani AW from soft sciences discipline (as opposed to the results from past research) relies on phrasal features more than the AW from hard sciences discipline.

(Affiliation)

[1]*PhD Candidate, Department of Applied Linguistics, Government College University, Faisalabad, Pakistan, Email: ahmad453@yandex.com*
[2]*Dean, Faculty of Arts and Social Sciences, Government College University, Faisalabad, Pakistan*
[3]*PhD Candidate, Department of Applied Linguistics, Government College University, Faisalabad, Pakistan*

INTRODUCTION

Academic Writing

AW is "a very general register, characterized as written language that has been carefully produced and edited, addressed to a large number of readers who are separated in time and space from the author, and with the primary communicative purpose of presenting information about some topic" (Biber & Conrad, 2009, p. 32). It is one of the widely researched register of English. The choice for this attention is caused by a number of reasons. The first reason is that AW is a familiarized register. It is admired because it is the thing the narcissist researchers do the best therefore they "find it to be inherently important and worthwhile" (Biber & Gray, 2016, p. 67). Secondly, AW is emphasized due to its importance in university education mainly because "it is the primary register that students must control for academic success" (Biber & Gray, 2016, p. 67). Furthermore, AW poses challenges at the advanced level of education. That is why the researchers devote significant energies to the study of AW to find ways to help the teachers involved in the teaching of AW convolutions to the novice writers. Lastly and most importantly, AW "is dramatically different linguistically from spoken" and "other written registers". "This linguistic difference" makes AW not only "challenging for students" but also "worthy of investigation" (Biber & Gray, 2016, p. 67). This research considers AW as a familiarized register being important at university level education, and thus worthy of investigation to report linguistic differences that pose challenges to the students as well as teachers.

Features Characterizing Academic Writing

The development of AW skills is considered essential at the university level (Zu, 2004). Therefore, when the students move (from college) to the university level, they are expected to move (from more general AW tasks) to more discipline-specific and specialized AW tasks (Nesi & Gardner, 2012). The reason (behind it) is that the advanced level AW "is widely recognized as an elaborated form of discourse that is grammatically complex" (Staples, Egbert, Biber & Gray, 2016, p. 2). Some of the previous AW research experts (Carter & McCarthy, 2006; Huddleston, 1984) associated grammatical

(Affiliation)

[1]*PhD Candidate, Department of Applied Linguistics, Government College University, Faisalabad, Pakistan, Email: ahmad453@yandex.com*
[2]*Dean, Faculty of Arts and Social Sciences, Government College University, Faisalabad, Pakistan*
[3]*PhD Candidate, Department of Applied Linguistics, Government College University, Faisalabad, Pakistan*

complexity with embedded clauses while others (Brown & Yule, 1983; Chafe, 1982; Hughes, 2005) related it to the use of subordinate clauses. This practice was challenged by another group of AW experts (Biber, 2006; Biber & Gray, 2016; Biber, Johansson, Leech, Conrad, & Finegan, 1999) proving that AW complexity is caused by the use of long noun phrases comprising several noun modifiers e.g. head nouns, pre-modifiers, and post-modifiers. Biber et al. (1999) particularly provided empirical evidence to establish that dependent clauses (that were previously thought to characterize AW) characterize spoken discourse, and noun phrases characterize AW.

Thus, building on the group of experts (Biber, 2006; Biber & Gray, 2016; Biber, Gray & Poonpon, 2011; Biber et al., 1999) this research considers clausal and phrasal features as characteristics of conversation and AW respectively. The reason (behind it) is that conversation is situational (i.e. produced in situational contexts) whereas AW is autonomous, decontextualized, and explicit (DeVito, 1966; Johns, 1997; Kay, 1977; Olson, 1977). Biber and Gray (2016) explain this difference more explicitly. For them (Biber & Gray, 2016, p. 68) conversation: "is highly interactive"; "is produced in real time" without any "pre-planning"; hearer and speaker both are present "in the same situation"; and "is addressed to specific individual participants". In contrast, AW is: addressed to a large number of readers who usually do not interact with the writer; contained with "a high degree of shared background knowledge" meant "for specialists in an academic discipline" (Biber & Gray, 2016, p. 68); and carefully planned and edited (Biber & Gray, 2016; Staples, et al., 2016). Therefore, this research aims to study the AW from hard, and soft sciences disciplines (see academic disciplines in literature review) in order to see disciplinary traditions in the use of lexicogrammatical features (see Table 2).

Research Question

- Which lexicogrammatical features characterize Pakistani AW from hard and soft sciences disciplines?

LITERATURE REVIEW

Research (Biber, 2006; Biber & Gray, 2016; Egbert, 2015; Gray, 2015) reported variations in the use of lexicogrammatical features (particularly complexity features) across different parameters e.g. AW discipline, AW genre, and AW register. Such as research articles from science disciplines contained phrasal (complexity) features more than those from humanities, and social sciences disciplines. Such differences (i.e. related to discipline, genre and register) can be discussed in relation with "the

communicative characteristics of these varieties" (Staples, et al., 2016, p. 3). For example, "persuasion" according to Hyland (2008, p. 16) is found to be "more explicitly interpretative and less empiricist" in humanities. Therefore, academic writers from the field of humanities make less frequent use of phrasal features. It implies that any debate on AW complexity has "to consider disciplinary and genre differences" (Staples, et al., 2016, p. 4). Therefore, this research aims to study the disciplinary differences underlying AW in the use of different types of lexicogrammatical features (see Table 2). Previously, clausal features were thought to cause AW complexity. This thought was based on Hunt's (1965) idea i.e. T-Unit (clausal subordination) characterized AW, and caused complexity in it (AW). Following this notion, later research (e.g. Beers & Nagy, 2011; Brown & Yule, 1983; Kroll, 1977; Larsen-Freeman, 2006; Ortega, 2003; Wolfe-Quintero, Inagaki & Kim, 1998 considering clause/T-unit as complexity measure) extensively studied T-Unit as an indicator of AW complexity. However, this practice was challenged in the later research (Bardovi-Harlig, 1992; Biber, Gray & Poonpon, 2011, 2013; Biber, Gray & Staples, 2014; Lu, 2011; Rimmer, 2006; Taguchi, Crawford & Wetzel, 2013), and phrasal features were empirically proved (in Biber, Gray & Poonpon, 2011, 2013; Biber, Gray & Staples, 2014) to cause AW complexity. Therefore, there had been "a growing call for the inclusion of phrasal features along with the more traditional clausal features" Staples et al. (2016, p. 4) in the research (e.g. Biber, 1988; Biber & Conrad, 2009; Biber & Gray, 2010; Biber et al., 2011, 2013; Norris & Ortega, 2009; Ravid & Berman, 2010) on AW development for L1 and L2 writers. Therefore, this research is aimed to study clausal features with an inclusion of phrasal, and intermediate features (see Table 2) to see which of them characterizes Pakistani AW the most.

Academic Writing Materials

There are different types of AW materials e.g. abstracts, book reviews, essays, literature reviews, research proposals, dissertations, and so on (see Smith, 2020). Dissertation (the focus of this research) is a written piece of text, typically amounting to 150-300 pages (Swales, 2004). It comprises a document (presenting research and findings) that is submitted in support of candidature for professional qualification or an academic degree (ISO, 1986). The main reason behind the focus of this research on dissertation is that dissertation is the neglected (Starfield & Ravelli, 2006), under-studied, under-theorized, and under-taught text of research (Kamler & Thomson, 2006; Lundell & Beach, 2002). However, dissertation has undergone a major development in the last decade i.e. availability in electronic version, usually in PDF format. This development has now made it easier to access

authentic and representative texts than the past times (Thompson, 2012).

Furthermore, dissertation has been found to be a difficult and challenging piece of writing. The size of a dissertation text, complex task of planning, synthesizing of one's reading, and sustaining an extended as well as coherent argument are the main factors that pose challenge to the dissertation writers. This challenge gets even more difficult particularly when the language (in which the thesis is to be written) is not one's mother language (Thompson, 2012). Leki, Cumming and Silva (2008) have also observed that post graduate student writers of second language face threats to their sense of identity, and linguistic difficulties in an effort to adjust themselves to the environment where their disciplinary expertise is neither easily recognized nor expressed. Paré, Starke-Meyerring and McAlpine (2009) have also added that "the linguistic and rhetorical complexities of the dissertation are simply inexpressible for most academics". Keeping it in view, this research has been aimed to explore lexicogrammatical features (see Table 2) in the doctoral dissertations in order to add to the existing research, and invite the interest of future research towards this vital area that has been found to be "a kind of present absence in the landscape of doctoral education. It was something that everyone worried about, but about which there was too little systematic debate and discussion" due to the "relative scarcity of well-theorized material about doctoral supervision and writing" (Kamler & Thomson, 2006).

Academic Disciplines

Academic disciplines are examined from at least five perspectives e.g. anthropological, historical, management, philosophical, and sociological (Krishanan, 2009) that (perspectives) differ from each other due to the emphasis they put on cultural practices, development of disciplines, disciplinary division of knowledge in relation to education and market demands, historical conditions, and nature and theory of knowledge (Abbot, 2001). This research employs philosophical perspective owing to the applied, and pure linguistic research choice. In fact, conventional division of knowledge is well rooted in philosophical perspective (Beecher & Trowler, 2001), and it is the only perspective that classifies knowledge considering epistemological features of a discipline (Russel, 2002).

Research (Biber, 2006; Hyland, 2006, 2010) on linguistic variation across academic disciplines reports the disciplines to rely on lexicogrammatical features to realize communicative purposes. Thus, it means that this impression is well established for the reason "disciplines differ in their epistemological beliefs, research practices, and knowledge structures" (Gray, 2015, p. 1). In fact, linguistic variations

(across disciplines) emerge as a result of expectations of discourse community members (Hyland, 1998). Furthermore, linguistic feature variations are caused owing to the intrinsic differences lying between knowledge construction, and research practice disciplines (Charles, 2003). It implies that the different disciplines characterize different lexicogrammatical features. Therefore, this research aims to see which lexicogrammatical features characterize Pakistani AW (which, according to Azher, Ali and Mahmood (2021) is an interesting area for the investigation of internal and external variations) from hard and soft sciences disciplines. This disciplinary classification is based on two continua of pure/applied (soft/hard) sciences (see Becher & Trowler, 2001). Hence, four disciplines i.e. linguistics and history (soft sciences), and biology and physics (hard sciences) are chosen for this research.

METHODOLOGY

This is a corpus-based descriptive research. It is aimed to describe the salient features characterizing Pakistani AW. Corpus for this research comprises doctoral dissertations from two disciplines i.e. hard and soft sciences. The dissertations have been retrieved from an online source i.e. Pakistan Research Repository (PRR). It (PRR) is hosted by the Higher Education Commission of Pakistan, and contains a large number of dissertations written by Pakistani master and doctoral students. The subjects, and number of dissertations from the said disciplines is given in Table 1.

Table 1. Subject, Discipline and Number of Dissertations Used to Develop the Corpus of This Research

Sr.	Discipline	Subjects	Number of Dissertations	Number of Words
1	Hard Sciences	Biology	10	179080
		Physics	10	109361
2	Soft Sciences	History	10	27753
		Linguistics	10	489827
		TOTAL	**40**	**806021**

Source: Authors

Corpus development process involved simple procedural steps. In the first step, PDF versions of the doctoral dissertations were retrieved from PRR. In the second step, the PDF versions were converted into MS word format using an online available tool i.e. PDF converter. Cleansing process was applied

in the third step through which tables, figures, headings, references, and preliminary pages were removed from the data in MS word format. In the final step, the data was saved in Notepad files for processing in corpus tools.

As far as the corpus analysis was concerned, it was also carried through a number of simple procedural steps. First of all, the corpus was tagged with the help of MAT, and TagAnt Taggers. The tagged corpus was then processed in AntConc for analysis purpose that (AntConc processing) provided the results (see Tables 3 and 4) of different lexicogrammatical features in the corpus in the form of frequencies. The said frequencies were used to decide about the characteristic features of Pakistani AW.

The lexicogrammatical features (used in this research) were adopted from Staples et al. (2016) which comprised different types of clausal, intermediate, and phrasal features (Table 2). These features were searched in the corpus with the help of formulas devised in Ahmad (2022), a doctoral dissertation on Pakistani AW. For example, a formula (*_NN *_NN) was applied to find noun pre-modifiers in the corpus.

Table 2. Lexicogrammatical Features Employed in This Research

Sr.	Features	Types
1	Clausal	Finite adverbial clauses
		WH complement clauses
		Verb + *that*-clauses
		Clausal coordinating conjunctions
2	Phrasal	Nouns
		Attributive adjectives
		Premodifying nouns
		Nominalizations
		of genitives
		Prepositional phrases
3	Intermediate	Adverbs
		Linking adverbials
		Extraposed Adjective + *that* clauses
		Noun + *that*-clauses
		WH relative clauses
		That relative clauses

		Verb + *to*-clauses
		Desire verb + *to*-clauses
		Raising structures and extraposed adjective + *to*-clauses
		Noun + *to*-clauses
		Verb + *ing*-clauses
		Passive voice verbs
		Passive nonfinite
		relative clauses

Source: Staples et al. (2016)

ANALYSIS AND DISCUSSION

This research was aimed to find the lexicogrammatical features characterizing Pakistani AW from hard and soft sciences disciplines. The results (Table 3), in this regard, showed different use of clausal, phrasal, and intermediate features in hard and soft sciences disciplines. For example, the writers from soft sciences used different types of clausal features in different frequencies i.e. finite adverbial clauses (408 times), WH complement clauses (45 times), verb + *that*-clauses (66 times), and clausal coordinating conjunctions (1672 times). On the other hand, the writers from hard sciences also used different types of the same features in different frequencies i.e. finite adverbial clauses (58 times), WH complement clauses (17 times), verb + *that*-clauses (43 times), and clausal coordinating conjunctions (464 times). These results revealed the use of clausal features more in soft sciences discipline than in hard sciences discipline reflecting the use of clausal coordinating conjunctions, and *WH* complement clauses in maximum, and minimum frequencies respectively. It means that clausal features characterize soft sciences AW more than hard sciences AW.

Similarly, the results (Table 3) showed the use of different types of phrasal features in different frequencies in both soft, and hard sciences disciplines. The use of nouns, attributive adjectives, premodifying nouns, nominalizations, *of* genitives, and prepositional phrases was found to be 165705, 29754, 23592, 25388, 10499, and 9810 times respectively in soft sciences discipline. In comparison, the use of similar features i.e. nouns, attributive adjectives, premodifying nouns, nominalizations, *of* genitives, and prepositional phrases was observed to be 93218, 17256, 18128, 12518, 5853, and 4607 times respectively. In this way, soft sciences AW was found to contain phrasal features more than hard sciences AW. In addition, the use of nouns was high, and the use of prepositional phrases was low

across both disciplines. Therefore, phrasal features were found to characterize soft sciences AW more than hard sciences AW.

As far as the use of intermediate features was concerned, it was observed that the AW from soft sciences disciplines contained (Table 3) adverbs 1089 times, linking adverbials 3416 times, extraposed adjective + *that* clauses 26 times, noun + *that*-clauses 536 times, *WH* relative clause 204 times, *that* relative clauses 2014 times, verb + *to*-clauses 1347 times, desire verb + *to*-clauses 0 times, raising structures and extraposed adjective + *to*-clauses 4 times, noun + *to*-clauses 2506 times, verb + *ing*-clauses 149 times, passive voice verbs 13 times, and passive nonfinite relative clauses 941 times. Similarly, AW from hard sciences discipline used different types of intermediate features with different frequencies i.e. adverbs 339 times, linking adverbials 1661 times, extraposed adjective + *that* clauses 11 times, noun + *that*-clauses 101 times, *WH* relative clauses 77 times, *that* relative clauses 492 times, verb + *to*-clauses 239 times, desire verb + *to*-clauses 0 times, raising structures and extraposed adjective + *to*-clauses 0 times, noun + *to*-clauses 661 times, verb + *ing*-clauses 126 times, passive voice verbs 3 times, and passive nonfinite relative clauses 735 times. In this way, the use of intermediate features was found to be more in soft sciences AW as compared to that of hard sciences AW. Therefore, intermediate features were found to characterize soft sciences AW more than hard sciences AW. In addition, the use of linking adverbials was found to be high, and the use of raising structures and extraposed adjective + *to*-clauses was found to be low across both disciplines.

Table 3. Frequencies of Different Types of Clausal, Phrasal and Intermediate Features

DISCIPLINE	SOFT SCINCES			HARD SCIENCES		
Clausal Features	**Frequencies in History**	**Frequencies in Linguistics**	**Total**	**Frequencies in Biology**	**Frequencies in Physics**	**Total**
Finite adverbial clauses	6	402	408	24	34	58
WH complement clauses	0	45	45	7	10	17
Verb + *that*-clauses	0	66	66	14	29	43

Clausal coordinating conjunctions	94	1578	1672	248	216	464
Phrasal Features						
Nouns	9045	156660	165705	50941	42277	93218
Attributive adjectives	1696	28058	29754	8560	8696	17256
Premodifying nouns	1732	21860	23592	10365	7763	18128
Nominalizations	1230	24158	25388	5931	6587	12518
of genitives	688	9811	10499	3277	2576	5853
Prepositional phrases	707	9103	9810	2141	2466	4607
Intermediate Features						
Adverbs	54	1035	1089	162	177	339
Linking adverbials	129	3287	3416	634	1027	1661
Extraposed Adjective + *that* clauses	0	26	26	2	9	11
Noun + *that*-clauses	10	526	536	41	60	101
WH relative clauses	5	199	204	28	49	77
That relative clauses	48	1966	2014	258	234	492
Verb + *to*-clauses	10	1337	1347	93	146	239
Desire verb + *to*-clauses	0	0	0	0	0	0

Raising structures and extraposed adjective + *to-* clauses	1	3	4	0	0	0
Noun + *to-* clauses	87	2419	2506	302	359	661
Verb + *ing-* clauses	3	146	149	66	60	126
Passive voice verbs	1	12	13	1	2	3
Passive nonfinite relative clauses	43	898	941	364	371	735

Source: Authors

Overall, the results (Table 4) showed that the use of phrasal, intermediate, and clausal features was the first highest, second highest, and third highest respectively in both disciplines. Whereas, AW from soft sciences discipline contained clausal features 2191 times, phrasal features 264748 times, and intermediate features 12245 times. In contrast, the overall use of phrasal, clausal, and intermediate features was observed to be 582, 151580, and 4445 times respectively. These results indicated the use clausal, phrasal, and intermediate features being more in soft sciences AW than in hard sciences AW. Therefore, phrasal, intermediate, and clausal features were found to characterize AW from soft sciences discipline more than the AW from hard science discipline.

Table 4. Comparison of Clausal, Phrasal and Intermediate Features in Hard and Soft Sciences Disciplines

DISCIPLINE	SOFT SCIENCES			HARD SCIENCES		
Features	History	Linguistics	Total	Biology	Physics	Total
Clausal	100	2091	2191	293	289	582

Phrasal	15098	249650	264748	81215	70365	151580
Intermediate	391	11854	12245	1951	2494	4445

Source: Authors

Comparing the results (as shown in Tables 3 and 4) with those of as described in Staples et al. (2016) it is stated that Staples et al. (2016) report an increased use of the phrasal features in the advanced level AW particularly, noun+ *of* phrases, nouns, nominalizations, and attributive adjectives have been reported with an increased use in social sciences. Similarly, use of *of* genitives, and prepositional phrases has been reported with an increased use in arts and humanities. In contrast, premodifying nouns have been reported with an increased use in life, and physical sciences disciplines.

The results of this research (Tables 3 and 4) also show phrasal features in the highest use in both hard, and soft sciences disciplines. However, the use of all phrasal features (see Tables 2 and 3) is also higher in soft sciences discipline as compared to the use of phrasal features in hard sciences discipline (this research treats arts and humanities, and social sciences as soft sciences, and life sciences, and physical sciences as hard sciences). So, there is no noteworthy difference in the use of phrasal features in the Pakistani AW from hard, and soft sciences disciplines.

As far as the use of clausal features is concerned, Staples et al. (2016) report clausal features with a decreased use across disciplines. Similar trend can also be observed from the results of this research (Tables 3 and 4) i.e. Pakistani AW across both disciplines (i.e. hard and soft sciences) has been found to contain clausal features in low frequencies as compared to the frequencies of phrasal features. Furthermore, Staples et al. (2016) report more frequent use of finite adverbial clauses in arts and humanities, and social sciences disciplines, and less frequent use in life and physical sciences disciplines. In contrast, the use of finite adverbial clauses is found (in this research) to be in the second highest frequency (the first being clausal coordinating conjunctions) in soft sciences discipline. The same trend (see Table 3) can be seen in hard sciences discipline. In addition, Staples et al. (2016) report the frequent use of *WH* clauses, and noun + *that* complement clauses in arts and humanities discipline. These results are averse to the results of this research (see Table 3).

Taken together, Staples et al. (2016) report a fair decrease, and increase in the use of clausal, and phrasal features (respectively) across disciplines. Furthermore, clausal features have been reported (in Staples et al., 2016) to play an important role in arts and humanities, and social sciences disciplines than in life sciences, and physical sciences disciplines. In contrast, phrasal features have been reported

(in Staples et al., 2016) in an increased use across disciplines. The results of this research (Tables 3 and 4) also show the similar trend i.e. there is a high use of phrasal, and low use of clausal features across disciplines, and clausal, phrasal and intermediate features (in contrast with clausal features playing an important role in arts and humanities) all play an important role in soft sciences discipline. The results (Tables 3 and 4) of this research show that Pakistani AW from hard, and soft sciences disciplines characterizes phrasal features. These results align with the results of other studies e.g. Biber and Gray (2016), Gray (2015), and Staples et al. (2016). However, the results of this (present) research are also different from the results of the said research (i.e. Biber & Gray, 2016; Gray, 2015; Staples et al., 2016) in the sense that the said research unanimously reports AW from hard sciences disciplines (life sciences, and physical sciences) relying more on phrasal features than AW from soft sciences disciplines (arts and humanities, and social sciences). Thus, the results of this (present) research show Pakistani AW from soft sciences discipline relying more on phrasal features more than AW from hard sciences discipline.

CONCLUSION

This research was aimed to find the salient features characterizing Pakistani AW from hard, and soft sciences disciplines. Results revealed the use of clausal features more in AW from soft sciences discipline than in the AW from hard sciences discipline reflecting the use of clausal coordinating conjunctions, and *WH* complement clauses in maximum, and minimum frequencies respectively. It meant that clausal features characterize AW from soft sciences more than the AW from hard sciences. Similarly, the results revealed AW from soft sciences containing phrasal features more than the AW from hard sciences, and the use of nouns was high, and the use of prepositional phrases was low in AW across both disciplines. Therefore, phrasal features were also found to characterize AW from soft sciences more than the AW from hard sciences. In the same way, intermediate features were found to characterize AW from soft sciences more than the AW from hard sciences, and the use of linking adverbials was found to be high, and the use of raising structures and extraposed adjective + *to*-clauses was found to be low in the AW across both disciplines. These results indicated the use of clausal, phrasal, and intermediate features being more in the AW from soft sciences than in the AW from hard sciences. Therefore, phrasal, intermediate, and clausal features were found to characterize AW from soft sciences discipline more than AW from hard sciences discipline. Furthermore, the use of phrasal, intermediate, and clausal features was found to contain in the first, second, and third highest frequent

use. Therefore, phrasal features were concluded to characterize Pakistani AW. In addition, phrasal features were found to characterize AW from soft sciences discipline more than the AW from hard sciences discipline. Therefore, it was concluded that Pakistani AW from soft sciences discipline (as opposed to the results of past research) relies on phrasal features more than the AW from hard sciences discipline.

Note: This article has been extracted partially from a doctoral dissertation titled: "Phrasal complexity in Pakistani academic writing: A corpus-based comparative study of doctoral dissertations across disciplines".

REFERENCES

Abbott, A. (2001). *Chaos of disciplines.* Chicago and London: University of Chicago Press.

Ahmad, M. (2022). *Phrasal complexity in Pakistani academic writing: A corpus-based comparative study of doctoral dissertations across disciplines* (Doctoral Dissertation). Department of Applied Linguistics, Government College University, Faisalabad, Pakistan.

Azher, M., Ali, S., & Mahmood, M. A. (2021). Linguistic variation across disciplines in Pakistani academic writing: A multidimensional analysis. *Jahan-e-Tahqeeq*, *4*(1), 43–67.

Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, *11*(1), 17–34.

Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories*. Buckingham: The Society for Research into Higher Education.

Beers, S. F., & Nagy, W. E. (2011). Writing development in four genres from grades three to seven: Syntactic complexity and genre differentiation. *Reading and Writing*, *24*(2), 183–202.

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

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.

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

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

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, *9*(1), 2–20.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, *45*(1), 5–35.

Biber, D., Gray, B., & Poonpon, K. (2013). Pay attention to the phrasal structures: Going beyond T-units—A response to WeiWei Yang. *TESOL Quarterly*, *47*(1), 192–201.

Biber, D., Gray, B., & Staples, S. (2014). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, *37*(5), 639–668.

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

Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English*. Cambridge: Cambridge University Press.

Chafe, W. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and Written Language: Exploring Orality and Literacy* (pp. 35–54). Norwood: Ablex.

Charles, M. (2003). 'This mystery…': A corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes, 2*(4), 313–326.

DeVito, J. A. (1966). Psychogrammatical factors in oral and written discourse by skilled communicators. *Speech Monographs, 33*(1), 73–76.

Egbert, J. (2015). Publication type and discipline variation in published academic writing: Investigating statistical interaction in corpus data. *International Journal of Corpus Linguistics*, *20*(1), 1–29.

Gray, B. (2015). *Linguistic variation in research articles: When discipline tells only part of the story*. Amsterdam: John Benjamins.

Huddleston, R. (1984). *Introduction to the grammar of English*. Cambridge: Cambridge University Press.

Hughes, R. (2005). *English in speech and writing: Investigating language and literature*. New York: Routledge.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. Champaign: National Council of Teachers of English.

Hyland, K. (1998). *Hedging in scientific research articles.* Amsterdam: John Benjamins Publishing

Company.

Hyland, K. (2006). *English for academic purposes: An advanced resource book.* Routledge.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4–21.

Hyland, K. (2010). Constructing proximity: Relating to readers in popular and professional science. *Journal of English for Academic Purposes, 9*(2), 116–127.

ISO. (1986). *Documentation—Presentation of theses and similar documents*. Geneva: International Organization for Standardization.

Johns, A. (1997). *Text, role, and context: Developing academic literacies*. Cambridge: Cambridge University Press.

Kamler, B., & Thomson, P. (2006). *Helping doctoral students write: Pedagogies for supervision*. London: Routledge.

Kay, P. (1977). Language evolution and speech style. In B. Blount & M. Sanches (Eds.), *Sociocultural Dimensions of Language Change* (pp. 21–33). New York: Academic Press.

Krishnan, A. (2009). *What are academic disciplines? Some observations on the disciplinarity vs. interdisciplinarity debate.* UK: National Centre for Research Methods.

Kroll, B. (1977). Ways communicators encode propositions in spoken and written English: A look at subordination and coordination. In E. O. Keenan & T. Bennett (Eds.), *Discourse across Time and Space* (SCOPIL no. 5) (pp. 69–108). Los Angeles: University of Southern California.

Larsen–Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27*(4), 590–619.

Leki, I., Cumming, A., & Silva, T. (2008). *A synthesis of research on second language writing in English*. New York: Routledge.

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writer's language development. *TESOL Quarterly*, *45*(1), 36–61.

Lundell, D. B., & Beach, R. (2002). Dissertation writers' negotiations with competing activity systems. In C. Bazerman & D. Russell (Eds.), *Writing Selves/Writing Societies: Research from Activity Perspectives* (pp. 483–514). WAC Clearinghouse and Mind, Culture, and Activity: Colorado State University.

Nesi, H., & Gardner, S. (2012). *Genres across the disciplines: Student writing in higher education*. Cambridge: Cambridge University Press.

Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in SLA: The case of complexity. *Applied Linguistics*, *30*(4), 555–578.

Olson, D. (1977). From utterance to text: The bias of language in speech and writing. *Harvard Educational Review*, *47*(3), 257–281.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, *24*(4), 492–518.

Paré, A., Starke-Meyeering, D., & McAlpine, L. (2009). The dissertation as multi-genre: Many readers, many readings. In C. Bazerman, A. Bonini & D. Figueiredo (Eds.), *Genre in a Changing World* (pp. 179–194). Colorado and Indiana: The WAC Clearing House and Parlor Press.

Ravid, D., & Berman, R. (2010). Developing noun phrase complexity at school age: A text-embedded cross-linguistic analysis. *First Language*, *30*(3), 3–26.

Rimmer, W. (2006). Measuring grammatical complexity: The gordian knot. *Language Testing, 23*(4), 497–519.

Russell, D. R. (2002). *Writing in the academic disciplines: A curricular history.* Carbondale: Southern Illinois University Press.

Smith, S. (2020). *Academic writing*. Retrieved on December 11, 2022 from https://www.eapfoundation.com/writing/what/.

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), 1–35.

Starfield, S., & Ravelli, L. J. (2006). "The writing of this thesis was a process that I could not explore with the positivistic detachment of the classical sociologist": Self and structure in new humanities research theses. *Journal of English for Academic Purposes*, *5*(3), 222–243.

Swales, J. M. (2004). *Research genres: Explorations and applications*. Cambridge: Cambridge University Press.

Taguchi, N., Crawford, W., & Wetzel, D. Z. (2013). What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program. *TESOL Quarterly*, *47*(2), 420–430.

Thompson, P. (2012). Thesis and dissertation writing. In B. Paltridge & Sue Starfield (Eds.), *The Handbook of English for Specific Purposes* (pp. 283–299). UK and USA: John Wiley & Sons.

Wolfe–Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity.* Honolulu: University of Hawaii Press.

Zhu, W. (2004). Faculty views on the importance of writing, the nature of academic writing, and teaching and responding to writing in the disciplines. *Journal of Second Language Writing*, *13*(1), 29–48.