

## PalArch's Journal of Archaeology of Egypt / Egyptology

### EXPLORATION OF NEW DIMENSIONS IN ARGUMENTATIVE ESSAY WRITING IN A CORPUS OF WORLD ENGLISHES

*Dr Maimoona Abdulaziz<sup>1</sup>, Dr Muhammad Asim Mahmood<sup>2</sup>*

<sup>1</sup>Assistant Professor Department of English Riphah International University Faisalabad

<sup>2</sup>Professor Department of Applied Linguistics Government College University Faisalabad

Email: [1abdulaziz.maimoona@gmail.com](mailto:1abdulaziz.maimoona@gmail.com) [2asimrai@gmail.com](mailto:2asimrai@gmail.com)

**Dr Maimoona Abdulaziz, Dr Muhammad Asim Mahmood. Exploration Of New Dimensions in Argumentative Essay Writing in A Corpus of World Englishes -- PalArch's Journal of Archaeology of Egypt/Egyptology 19(1), 1488-1502. ISSN 1567-214x**

**Key Words: Multidimensional Analysis, Argumentative Essays, Learner Writing, World Englishes, Corpus Analysis**

#### ABSTRACT

This study is an exploration of new functional dimensions in argumentative essay writing of learners and employs Multidimensional Analysis technique for the said purpose. A corpus encapsulating all three divisions of World Englishes has been analyzed comprising over 2500 student essays written by ENL, ESL, EFL learners. Three new dimensions surfaced as a result of the analysis, based on 67 prominent linguistic features. After a thorough review of relevant literature and expert opinion, the three resultant dimensions have been interpreted as Abstract Information vs. Concrete Description, Contextualized Interactive Elaboration vs. Incoherent Informational Discourse, and Author's Personal Stance and Synthesis. This paper is beneficial for future researchers who intend to compare learner writing of different varieties of World Englishes. It is also significant from a pedagogical perspective, as the results may be used as benchmarks for assessment of argumentative essay writing.

#### INTRODUCTION

Learner writing is recognized as an insignia of written language development as it depicts current language situation (Granger, 2002). Over the years, it has generated considerable interest for the researchers in terms of language pedagogy and development. Initially, the focus was on identifying lexical and syntactical errors (e.g. Sarfraz, 2011; Sajid & Siddiqui, 2015), exploring cultural and religious aspects (e.g. Mahboob, 2009), or looking into individual linguistic features (e.g. Gardezi & Nesi, 2009), and a number of corpus-based researches have taken place in this regard. The major drawback of these researches is that

they do not account for a holistic view of language functionality, and mostly incorporate subjective analyses based on author's observation. Quite recently, however, the need for a deeper analysis has been acknowledged and attention has been shifted to exploring the functional aspects of learner language based on sets of co-occurring features that define them (e.g. Azher & Mahmood, 2016). The trend started off with genre-based studies (e.g. Imtiaz & Mahmood, 2014) and moved on towards recent approaches of register analysis, more specifically, multi-dimensional analysis as given in Biber (1988) (e.g. Abdulaziz et al. 2016). The analysis identifies and interprets sets of co-occurring linguistic features in terms of their shared communicative functions by using computer software tools like linguistic taggers and statistical tools like factor analysis. The methodology has gained immense popularity in the recent years owing to its objectivity, and empirical bent towards language analysis. It began as a corpus-based technique to find linguistic variation between spoken and written registers based on functional dimensions (e.g. Biber, 1988). However, several publications in recent years have appeared documenting multi-dimensional analyses of more specific registers or sub-registers, like academic essays (Friginal & Weigle, 2014), narrative and expository writing (Asencion-Delaney & Collentine, 2011) etc. The results obtained from these studies propose unique functional dimensions specific to their respective sub-registers.

### ***Problem Statement***

The genre of argumentative essay writing has been explored using different approaches, like genre analysis and error identification, while most of these studies are subjective in nature and focus on drawing comparisons between one or two ESL/EFL varieties and the native variety. However, to the author's best knowledge, no such work has been carried out as yet that focuses on identifying functional dimensions specific to argumentative essays.

### ***Aim and Scope***

This research aims to make one such effort as it incorporates the use of corpus-based multidimensional analysis (Biber, 1988) in order to explore functional and linguistic variation, based on quantifiable evidence of co-occurring patterns of language use or sets of co-occurring features. The results may be used for generating quantitative comparisons of different language varieties based on student writing in order to have a current view of their language development process.

## **LITERATURE REVIEW**

As is the case, human communication exhibits variation at different levels including morphological, syntactical, prosodic, pronunciation, discourse, and lexical levels. Ure (1982) argues that every speech community has its own system of registers depending on the range of everyday activities of its members. Their word choice, prosody, and grammar move according to their need and demand. Cultural differences may, therefore, contribute to register variation. According to Hymes (1984), it is next to impossible that a human being speaks the same language all the time without causing any variation in his/her register or style. Variation, therefore, is an innate characteristic of human

communication. This variation across registers has been of key interest to researchers who have made several approaches towards comparative study of registers from various perspectives.

### ***Investigating Linguistic Variation: From Register Analysis to MD Analysis***

Earlier studies on registers focused mainly on analysis of linguistic characteristics that are commonly found in a particular text variety combined with the analysis of its situation, leading on to an investigation of their communicative purposes. Biber and Conrad (2009) argue that linguistic features are always functional when considered from a register perspective:

...linguistic features tend to occur in a register because they are particularly well suited to the purposes and situational context of the register. Thus, the third component of any register description is the functional analysis. (p. 6)

Register analysis usually follows a three-step procedure; firstly, it observes distinguished situational characteristics of the communication, secondly it takes multiple texts from the target situation in order to identify frequently occurring features, and finally it interprets the relationship between these situational characteristics and prevalent linguistic features in purely functional terms (Biber & Conrad, 2009). The functional interpretation of the text, therefore, is based on a comparison of its situational and linguistic analyses.

Other perspectives include the genre perspective, which explains situational varieties by analysing characteristic language features of texts. These genre features, however, are not pervasive, and are delimited to conventional rather than functional view of language. Additionally, they might not occur frequently in a text. From the perspective of style, the variation is studied from aesthetic point of view which is not directly functional, and hence is not the concern here and will not be discussed.

These approaches towards register analysis have been traditional in that they are confined to identifying small sets of features in a limited set of data, which affects the comprehensiveness of its results. The need to compare entire domains of language use with respect to a comprehensive set of lexical and grammatical features, has led to the advent of what we may term as a technological advancement in the field of applied linguistics. It has given rise to a multidimensional approach developed for such large scale analyses and comparisons on several linguistic parameters in highly quantitative terms.

### ***Relevance of MD Approach for Learner Corpus Analysis***

The advanced analytical approach towards corpus is one step further in that it is multidimensional in nature and aims at identifying situational factors on the basis of frequently co-occurring linguistic characteristics in purely quantitative terms. Its ultimate goal is to achieve a comprehensive description of linguistic variation and use in a language (Biber & Conrad, 2009). The present study aims to employ this analysis technique in order to explore the underlying functionality of language in one particular register. Recently, researches on

learner corpus following this multidimensional approach are surfacing, based on learner corpora such as ICLE, BAWE, LINDSEI, etc. (as in Van Rooy, 2008; Van Rooy & Terblanche, 2009; Xiao, 2009; Xiao & Cao, 2013, Crossley et al., 2014; Egbert, 2014; 2015). The succinct yet meticulous review of the methodologies employed in the analyses of learner corpora world-wide spanning over the last decade not only underscores the relevance of MDA approach to learner corpus but is also indicative of the present need of assessing world-wide learner corpus in terms of functionality.

### ***Research Questions***

When it comes to applying a particular methodology to a set of data, an important aspect that needs to be covered is the contextualization of the results to theory and practice in its relevant domain. Based on this framework, this research focuses on answering one major and three minor questions:

- What are the underlying dimensions of linguistic variation in argumentative essays?
  - What are the frequently co-occurring linguistic features of learner corpus according to new MDA?
  - What are the new factors specific to argumentative essays?
  - What is the functional interpretation of the new factors?

## **METHODOLOGY**

### ***Research Design***

The corpus design for present study was based on Kachruvian division of World Englishes into three categories (circles), English for Native Learners (ENL), English as a Second Language (ESL), and English as a Foreign Language (EFL), the primary goal of which was to illustrate the unprecedented variability in English (Kachru, 1985). Braj Kachru's Three Circle Model sets out to demonstrate the types of varieties that have surfaced over the years with the spread of English. This study has been designed on the same lines; viz. a variety of ESL circle namely Pakistani English has been compared with the three Kachruvian categories in order to determine its present position on the Kachruvian cline. The corpus that has been sampled for this study comprises of argumentative essays written by students, as detailed below.

### ***The ICNALE Corpus***

The corpus used for the present research comprises of learner essays extracted from International Corpus Network of Asian Learners of English (ICNALE). It is a large-scale digitalized database of learner corpora and includes 1.8 million words of controlled L2 English speeches and essays. It is considered as a reliable database for conducting international contrastive language analyses. The corpus is available online publicly and is easily downloadable under the Creative Common License. It comprises of two major modules: The ICNALE-Spoken and the ICNALE-Written. The ICNALE-Written has been completed in 2013 and contains more than 1,300,000 tokens of essays written by almost 2800

learners from fifteen different countries. These countries have been categorized under three groups; viz. ENL, ESL, and EFL.

***Research Participants. The Participants of This Research Are Principally English***

language learners. Details regarding their age, country, gender, name of school, academic major, years of studying English, and English proficiency level have been provided by ICNALE in an Excel sheet. The L2 proficiency of ESL and EFL learners has been classified into four levels; A2, B1\_1, B1\_2, and B2+, based on standard L2 proficiency tests, similar to the ones given in the Common European Framework of Reference (CEFR).

**Population groups.** The corpus has been categorized based on Kachru's division of World Englishes into three circles/ categories; ENL, ESL, and EFL. **Criteria for sampling.** Data from levels A2 and B1\_1 could not be considered for the research because of relatively lower proficiency of the learners as determined by their TOEIC/TOEFL/VTLS scores. Including them would obstruct the objective of comparison, which was not to evaluate errors made by the learners but to gauge language functionality and use. The sample therefore consisted of the following number of essays/ text files for each population group:

<b>Table 1</b> Sampled Data for The Present Research			
<b>Population Groups</b>	<b>Sub-Groups</b>	<b>Sample Size</b>	<b>Total Sample Size</b>
<b>ENL</b>		200	200
<b>ESL</b>	Hong Kong	138	1090
	Pakistan	178	
	Philippines	374	
	Singapore	400	
<b>EFL</b>	China	236	1242
	Indonesia	172	
	Japan	134	
	Korea	328	
	Thailand	204	
	Taiwan	168	

**Delimitations.** The essays were delimited to two per learner in order to have variation in sampling. Further, the topics for essays were argumentative in nature so that maximum expression in writing might be generated from the learners.

***Data Analysis***

A factor solution for new MD analysis is calculated based on 180+ linguistic features on which the data had been tagged by Biber's tagger. They are further short-listed based on the following criteria (Biber, 2015):

- All the linguistic features relevant to learner language, essay writing, academic prose, and expository writing have been included, as identified in previous studies.
- Communalities and factor structure have been considered, and features that do not contribute to the analysis due to very low values (0.1-0.9) have been discarded.
- Same features under different names have been excluded, and so are most of the hierarchically related features.

Once the shortlisting of features is done, a principal factor analysis has been applied in order to extract the maximum amount of shared variance among the linguistic features for each factor. The eigenvalue scree plot of this analysis is given below:

**Figure 1** Scree plot for principal factor analysis showing eigenvalues of factors. The plot shows the eigenvalues on y-axis and the number of factors on x-axis. High acceleration may be seen in the first three factors, which demarcates the largest proportion of variance. The eigenvalues as presented in Fig. 1 are basically indices of the amount of this variation, as well as, the percentage of shared variance, accounted for by each factor. One of the reasons for running factor analysis on this data is to reduce the larger number of factors that would describe a complex concept to a few interpretable latent variables or factors that can explain the maximum amount of variability in the data. The optimal number of coordinates is eighteen, which have been reduced to six due to their small number (Gorsuch, 1983).

Next, these factors have been rotated using Promax rotation in order to make theoretical sense of the analysis that will facilitate the interpretation of the constructs underlying each factor (as recommended in Biber, 1988). The inter-correlation of the six factors is as follows:

**Table 2.** Inter-Factor Correlations of the Six Factors Identified by New MDA

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Factor 1	1.00					
Factor 2	0.21	1.00				
Factor 3	0.03	0.12	1.00			
Factor 4	0.06	0.19	0.39	1.00		
Factor 5	0.24	-0.01	0.10	0.17	1.00	
Factor 6	-0.09	-0.31	-0.29	-0.29	-0.02	1.00

*Note.* Mean item complexity is 2.3. Test of the hypothesis that 6 factors are sufficient. The degrees of freedom for the null model are 7260 and the objective function is 71.48. The degrees of freedom for the model are 6549 and the objective function is 54.51. The root mean square of the residuals (RMSR) is 0.04 and the df corrected root mean square of the residuals is 0.04. Fit based upon off diagonal values is 0.79.

Once the features that co-occur are identified and grouped together as factors, a further scrutiny has been made, based on the standard cut-off point or threshold as set by Biber (1988) to +/-0.30, also known as the minimum weight of the feature. Since these weights indicate the tendency of the features to co-occur with other features on the same factor, it is therefore necessary to set a minimum weight limit, or salient loading, for all such features so that less significant features may be eliminated.

#### *Factors Based on New MDA*

The following chart shows the six sets of co-occurring features or factors that have been identified as a result of rotated factor analysis. Linguistic features with salient loadings, both negative and positive, have been highlighted (in bold) for each factor.

**Table 3** Rotated Factor Pattern Matrix for the Six Factor Solution Identified by New MDA with Salient Factor Loadings

Linguistic Features	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Private_Vb	0.23	0.28	<b>0.56</b>	0.06	-0.03	-0.06
That_Del	0.06	0.27	<b>0.43</b>	-0.02	-0.08	-0.08
Contract	-0.15	<b>0.35</b>	0.02	0.09	-0.02	-0.16
Vb_Present	-0.22	<b>0.72</b>	0.21	-0.08	0.02	0.12
Do_Pro	0.01	<b>0.32</b>	-0.07	-0.1	-0.02	-0.08
Emphatic	0	0.11	-0.02	<b>0.4</b>	-0.06	-0.06
Sub_Conj_Caus	-0.08	<b>0.35</b>	-0.12	-0.07	0	-0.01
Amplifr	<b>-0.33</b>	0.1	0.09	0	-0.1	0.06
Prep	0.03	<b>-0.48</b>	-0.06	-0.01	-0.13	-0.02

Jj_Attr	0.09	<b>-0.52</b>	-0.08	-0.07	-0.11	-0.2
Vb_Past	-0.09	-0.09	0.06	0.04	0.04	<b>-0.37</b>
Vb_Public	-0.13	0.02	<b>0.37</b>	-0.15	0.02	-0.06
Infinitive	0.3	-0.01	0.14	0.02	<b>0.49</b>	0.15
Sub_Conj_Cond	-0.11	<b>0.37</b>	-0.04	0	-0.01	0.06
Mod_Necess	<b>-0.35</b>	0.28	0.08	-0.12	-0.1	0.16
Split_Aux	<b>-0.35</b>	-0.14	0.09	0.21	0	0.18
Conj_Advl	0.03	-0.14	0.19	0.09	-0.02	0.32
Passive_Short	<b>-0.56</b>	-0.14	0.02	-0.05	-0.03	0.11
Th_Vb	-0.15	0.06	<b>0.69</b>	-0.09	0.01	0.18
Vb_Have	0.28	<b>0.31</b>	-0.01	-0.04	-0.08	0.05
Vb_Progress	<b>-0.56</b>	0.01	0.09	0	-0.03	0.06
Fact_Vb_Other	<b>0.38</b>	-0.07	0.16	0.06	-0.02	-0.12
Nn_Common	<b>0.74</b>	<b>-0.38</b>	-0.2	-0.22	-0.18	-0.03
Nn_Premod	<b>0.82</b>	-0.16	-0.1	-0.08	-0.04	0.03
Th_Vb_Comm	-0.1	-0.12	<b>0.37</b>	-0.15	0.12	-0.02
Th_Vb_Att	-0.08	0.01	<b>0.33</b>	-0.01	0.11	0.22
Th_Vb_Fact	-0.04	0.05	<b>0.34</b>	-0.04	0.1	-0.01
Th_Vb_Likely	-0.01	0.24	<b>0.58</b>	-0.01	-0.08	0.07
To_Vb_Desire	-0.11	0.27	0.02	-0.06	<b>0.48</b>	-0.05
To_Vb_Effort	0.03	-0.08	0.02	0	<b>0.46</b>	0.09
Advl_Fact	-0.05	-0.02	-0.11	<b>0.8</b>	-0.05	0.04
Advl_Likely	-0.03	0.02	-0.02	<b>0.35</b>	-0.06	0.05
Th_Vb_Stance_All	-0.11	0.12	<b>0.95</b>	-0.14	0.1	0.15
To_Vb_Stance_All	-0.11	0.12	0.11	-0.1	<b>0.92</b>	0
To_Jj_Stance_All	0.16	0.02	0.1	0.1	0.02	<b>0.31</b>
Advl_Stance_All	-0.06	-0.07	-0.17	<b>1.03</b>	-0.09	0.13
Nn_Process	<b>0.51</b>	-0.17	-0.06	-0.01	-0.02	0.02
Nn_Cog	<b>0.48</b>	-0.14	0.2	0.01	-0.16	0.08
Nn_Abstact	<b>0.67</b>	-0.05	-0.03	-0.06	-0.09	0.03
Nn_Concrete	<b>-0.38</b>	-0.05	-0.06	-0.03	-0.04	-0.08
Nn_Quant	<b>0.89</b>	-0.01	-0.02	0	-0.05	0.08
Nn_Place	<b>-0.55</b>	0.16	-0.06	-0.12	-0.09	0.03
Nn_Group	<b>0.7</b>	-0.08	0.06	0.01	0	0.13
Vb_Act	<b>0.34</b>	0.15	-0.11	0.01	0.15	0
Vb_Mental	0.13	<b>0.39</b>	<b>0.5</b>	0.02	0.18	-0.07
Tt_Ratio	-0.15	<b>-0.61</b>	-0.01	0.19	0.06	<b>-0.33</b>
Word_Length	-0.21	<b>-0.59</b>	-0.04	-0.08	-0.04	0.12

Following the same methodology as Biber's (1988), another checkpoint has been placed on the redundant appearance of linguistic features on more than one factor. In case a particular linguistic feature possesses salient loading on more than one factor, its weight on each such factor is compared, and it is included in the factor in which it has the highest loading in terms of absolute magnitude, irrespective of the +/- sign. Three such features viz. Type-Token Ratio, Mental



Verbs, and Common Nouns, have been identified (Table 4.6). Factors having less than five linguistic features, including the positive and the negative features, have been removed, leaving three prominent factors behind.

## DISCUSSION

**Interpretation of factors as dimensions.** The co-occurring linguistic features on each dimension refer to certain functions that language performs, though interpreting them is partly subjective and based on intuition. The three resultant factors, mentioned above, have been interpreted using knowledge and insight taken from previous researches, as well as expert opinions taken from researchers relevant to the field.

In order to interpret Dimension 1, the most prominent dichotomy has to be considered among its positive and negative features, viz. abstract nouns and concrete nouns. Abstract nouns usually contain non-technical philosophic notions, while concrete nouns refer to more technical specialized content. The co-occurrence of abstract nouns with five other nominal categories and noun-noun sequences on the positive end indicate highly informational content or textual density. Compared with these nominal categories, there are only two categories related to verbs, which make the use of nouns in such texts more pronounced. On the negative side, more prominent is the co-occurrence of two categories of verbs, layered with two adverbial categories, which is highly indicative of a more explanatory way of writing. A similar functional interpretation of factors derived from New MDA may be seen in Biber's own work, as Biber and Gray (2010) analyze a corpus of academic research articles amounting to around 3 million words categorized under four general disciplines. Stereotypical notion regarding complexity, elaboration, and explicitness in academic register have been empirically testified, keeping in view results from the past researches. Some innovation may be viewed in terms of the **Structural Elaboration vs. Compression** cline having similar grammatical features such as finite complement clauses, non-finite complement clauses, finite adverbial clauses, finite relative clauses, non-finite relative clauses on its positive side, whereas features such as attributive adjectives, pre-modifiers, prepositional phrase as noun post-modifier, appositive noun phrase as noun post-modifier, and prepositional phrase as adverbial are on its negative side, denoting structural compression.

Modal verbs of necessity co-occurring with verbs and adverbs refer to author's opinion about some particular action that needs to take place. Additionally, concrete nouns and place nouns provide references to real life objects and places. Keeping in mind features of both the sides and their functionality, Dimension 1 may be labeled as "Abstract Information vs. Concrete Description". Researchers have named the more nominalized end of the cline as **Informational Academic Discourse** (Azher & Mehmood, 2016), yet in this case, both the ends present academic discourse. Here, more prevalent are the dichotomies abstract-concrete and information-description/elaboration (as in Biber, 1988; Egbert, 2014, etc). This functional distinction is foreseeable in the example extracts taken from original texts in the table below.

<b>Table 4</b> Interpretation of Factors as Dimensions Based on New Multi-Dimensional Analysis with Examples	
<b>Dimension 1: Abstract Information vs. Concrete Description</b>	
<b>Features with Positive Loadings</b>	<b>W_CHN_PTJ0_228_B2_0 (D1 score: 11.1)</b>
Factive Verbs in other contexts (0.38)	As is so often been <b>emphasized</b> , practical <b>experience</b> should be <b>applied</b> to theoretical <b>knowledge</b> we've <b>learned</b> in the <b>class</b> . Having a <b>part-time</b> is an ideal <b>method</b> to increase one's practical <b>skills</b> and expand social <b>network</b> during the <b>college</b> (...) Such <b>cases</b> should be <b>taken</b> into <b>consideration</b> since it <b>happens</b> all the <b>time</b> when there's a big <b>project</b> in the <b>work</b> for you to conduct. The <b>lack</b> of the <b>ability</b> of <b>time management</b> can <b>make</b> the <b>situation</b> even worse. In <b>conclusion</b> , whether a part-time <b>job</b> is important to a <b>student depends</b> on the <b>attribution</b> of the <b>job</b> as well as the <b>ability</b> of the <b>student</b> .
Common Nouns (0.74)	
Pre-modifying Nouns/ N-N Sequences (0.82)	
Process Nouns (0.51)	
Cognitive Nouns (0.48)	
Abstract Nouns (0.67)	
Nouns of Quantity (0.89)	
Group/ Institution Nouns (0.7)	
Activity Verbs (0.34)	
<b>Features with Negative Loadings</b>	<b>W_HKG_SMK0_013_B1_2 (D1 score: -12.7)</b>
Adverb / Qualifier – Amplifier (-0.33)	Moreover, <b>restaurant</b> is a public <b>area</b> that different kinds of <b>people</b> can go into. Therefore, I agree that smoking <b>should</b> be <b>completely banned</b> at all <b>restaurants</b> for the sake of the mental as well as physical health of the <b>public</b> (...) As the major component in the <b>cigarette</b> is <b>tar, nicotine</b> etc. All of them are harmful to <b>human body</b> (...) Even worse, the <b>chemicals</b> inside the <b>cigarette</b> can cause a lot of diseases.
Modal of Necessity (-0.35)	
Adverb within Auxiliary/ Splitting aux-verb (-0.35)	
Agentless Passive Verb (-0.56)	
Verb – Present Progressive (-0.56)	
Concrete Nouns (-0.38)	
Place Nouns (-0.55)	
<b>Dimension 2: Contextualized Interactive Elaboration vs. Incoherent Informational Discourse</b>	
<b>Features with Positive Loadings</b>	<b>W_ENS_PTJ0_050_XX_1 (D2 score: 13.2)</b>
Contraction (0.35)	Although I <b>want</b> to <b>have</b> a part-time job my mom will not let me <b>because</b> she thinks it will interfere with my studying. Many of my friends are lucky enough to <b>have</b> parents that <b>let</b> them <b>do</b> basically what they <b>want</b> to <b>do</b> in college <b>as long as</b> they <b>have</b> good grades. I <b>have</b> good grades, but this <b>does</b> not <b>seem</b> to be good enough for my mom. I <b>tell</b> her that I am <b>doing</b> fine and I <b>have</b> plenty of time, and I would <b>love</b> to make a little bit of money <b>so that</b> I could <b>go</b> out with my friends <b>if</b> I wanted to, but she simply <b>won't</b> hear it.
Verb (uninflected present, imperative and third person) (0.72)	
Pro-verb "do" (0.32)	
Subordinating Conjunction - Causative (0.35)	
Subordinating Conjunction – Conditional (0.37)	
Verb 'have' (0.31)	
<b>Features with Negative Loadings</b>	<b>W_HKG_PTJ0_001_B1_2 (D2 score: -11.2)</b>
Preposition (-0.48)	There are <b>different</b> pros and cons <b>for</b> applying <b>part-time</b> jobs, and there are <b>different</b> kinds of <b>such</b> jobs. The <b>most popular part-time</b> job is sales <b>in</b> restaurants and <b>junior technical</b> support <b>in</b> computer shops. <b>Some</b> students would apply <b>for freelance</b> jobs which require
Attributive Adjective (-0.52)	
Type-Token Ratio (-0.61)	
Average Word Length (-0.59)	

	<b>advanced skill</b> sets, such as taking photographs or writing <b>computer</b> programs.
<b>Dimension 3: Author's Personal Stance and Synthesis</b>	
<b>Features with Positive Loadings</b>	<b>W_ENS_SMK0_044_XX_1 (D3 score: 23.5)</b>
Private Verb (0.56)	I absolutely <b>hate</b> smoking and I <b>find</b> it very difficult to hang out with smokers. I <b>think that it is quite rude of them to smoke in places where nonsmokers are present, especially restaurants</b> . How are you supposed to <b>enjoy</b> a nice meal when your eyes are watering and your throat is burning from the secondhand smoke? How are you supposed to <b>respect</b> the group of people who don't <b>value</b> their own health highly enough as to make smart choices which can <b>affect</b> the longevity of their existence. If you <b>think</b> about these questions, you will <b>realize that smokers who choose to smoke have forfeited their right to have a valid opinion on the subject</b> . I do <b>agree that they should be able to do whatever they want to do, but I don't think it is fair to nonsmokers to allow them to harm their health</b> .
'That' Deletion (0.43)	
Public Verbs (0.37)	
'That' Complement Clause Controlled by Verb (0.69)	
'That' Complement Clause Controlled by a Communication Verb (0.37)	
'That' Complement Clause Controlled by Mental/Attitudinal Verb (0.33)	
'That' Complement Clause Controlled by Factive Verb (0.34)	
'That' Complement Clause Controlled by Verb of Likelihood (0.58)	
Sum Stance 'That' Complement Clauses Controlled by Verbs (0.95)	
Mental Verbs (0.5)	
<b>Features with Negative Loadings</b>	<b>W_PAK_SMK0_155_B1_2 (D3 score: -9.4)</b>
No negative features	Smoking is very bad habit. As a result of which environment is polluted badly. It should be banned everywhere while the restaurant is a place, where people come to eat and to refresh themselves. At these places, if smokers come, due to smoking other people who are eating there are disturbed. Smoke is not only harmful for smoker but also for other peoples living around him.

Dimension 2 is unique in its co-occurring scheme. The positive end has verbs and proverb *do* co-occurring with subordinating conjunctions (causative and conditional) and contractions. As has been discussed in case of Dimension 1, verbs (have, and uninflected present tense) and proverbs signify elaboration and discussion. Contractions demonstrate a lower level of formality and more interactivity. Subordinating conjunctions combined with verbs work together to provide reasoning and context to the argument. What develops as a functional aspect based on the co-occurrence of these features therefore is a text that elaborates and discusses in an interactive manner, providing reasons and causes for the author's stance or point of view. Functional interpretation of factors derived from New MDA may be seen in Biber's own work as well, as Biber and Gray (2010) analyse a corpus of academic research articles amounting to around 3 million words categorized under four general disciplines. Stereotypical notion regarding complexity, elaboration, and explicitness in academic register have been empirically testified, keeping in view results from the past researches. Some innovation may be viewed in terms of the **Structural Elaboration vs. Compression** cline having grammatical features such as finite complement

clauses, non-finite complement clauses, finite adverbial clauses, finite relative clauses, non-finite relative clauses on its positive side, whereas features such as attributive adjectives, pre-modifiers, prepositional phrase as noun post-modifier, appositive noun phrase as noun post-modifier, and prepositional phrase as adverbial are on its negative side, denoting structural compression.

In another study in the field of corpus stylistics, Egbert (2014) examines a corpus of University textbooks in order to search for stylistic variation. The research introduces five new dimensions, one of which was **Academic Involvement and Elaboration vs. Information Density**, having positive features such as *be* as main verb, present tense verbs, activity verbs, prediction modals, predicative adjectives, to- clauses controlled by adjectives, demonstrative pronouns, conditional subordinators, causative subordinators, core vocabulary, and negative features as word length, attributive adjectives, and nouns. Though not directly related to the present research, dimensions revealing stylistic aspects of writing may serve relevance to essay writing and hence needs to be examined.

The negative end has prepositions and attributive adjectives, co-occurring with a high type-token ratio and word length. Informational as it may sound due to the presence of adjectives and prepositions, it is devoid of either verbs or subordination. Vocabulary is in bulk, as is signified by a high type-token ratio, yet due to lack of verbs, the discourse is relatively incoherent. Put together the features of both positive and negative ends, three functional dichotomies are prevalent; viz. contextualization-incoherence, elaboration-informational, and interactivity-informational. Dimension 2 may therefore be interpreted as “Contextualized Interactive Elaboration vs. Incoherent Informational Discourse”. The dichotomy of presence-absence of this combination of subordinating conjunctions with adverbs has also been interpreted as contextualized-detached (Azher & Mehmood, 2016), yet for argument-based writing, contextualized-incoherent seems to serve the purpose.

Lastly and most importantly, Dimension 3 is what may be termed as the most relevant cline for evaluating argumentation in texts. It is distinctive in that it has no negative features. The positive end is characterized by co-occurrence of *that* complement clauses and a highly loaded feature of sum stance *that* complement clauses (0.95) alongside verbs and that deletion. Stance, as the word suggests, is indicative of the author’s act of deciding on the negative or positive aspect of an argument. Coupled up with verbs such as mental verbs, it presents thinking or opinions of the stance-taker. The use of *that*-clauses controlled by verbs foregrounds the functional aspect of an argumentative text as synthesizing the results or opinions. Egbert (2014) takes it as a stylistic cline called **Research Synthesis** based on a corpus of University textbooks. For argumentative texts, however, it is more relevant to take it as a synthesis of the argument being developed. Collectively, the co-occurring features may be termed as “Author’s Personal Stance and Synthesis”.

New dimensions relevant to Pakistani academic writing have been explored by analysing a corpus of 235 MPhil and PhD theses comprising of around 8.4 million words (Azher & Mehmood, 2016). The analysis has been made on five

new dimensions with their respective sets of co-occurring linguistic features, the last of which is **Personal/ Evaluative Stance vs. Technical Description** having positive features as attitudinal adjectives, attitudinal verbs, all adjectives, predicative attributive adjectives, predicative adjectives, suasive verbs, stance nouns, and negative features such as technical nouns, all definite articles, group nouns.

In the stream of MDA researches conducted on published academic texts, Egbert (2015) highlights five dimensions based on a corpus of six registers under three publication types (journal articles, university textbooks, and popular academic books). One of these is **Author-centered Stance**, with positive features as first person pronouns, infinitives, nouns of human, cognitive nouns, mental verbs, suasive verbs, communication verbs, and stance reflected by a combination of that-clauses controlled by verbs: non-factive, factive, and likelihood, and that- relative clauses controlled by stance nouns.

In terms of investigating new registers such as internet based registers, Biber and Egbert (2016) made an inclusive research sampling from myriads of registers from the entire searchable web in order to find linguistic variation among those registers. The nine new dimensions based on these twenty seven specific web registers includes dimension of **Literate Stance**, having stance nouns + prep phrases, cognitive nouns, stance nouns + comp clauses, and other stance nouns as its positive features.

As has been mentioned earlier, the three new dimensions have been functionally interpreted based on a thorough review of related literature. For the purpose of verification, however, expert opinion has been taken from Dr. Jesse Egbert in this regard.

## CONCLUSION

The main objective of this research was to answer the research questions set forth in the light of a comprehensive literature review. The main query was to investigate dimensions specific to argumentative writing based data taken from all spheres of World Englishes. The sub-questions were directed towards highlighting frequently co-occurring linguistic features of learner corpus according to new MDA, as well as those specified to argumentative essays, as well as how this co-occurrence may be interpreted. The study has successfully given a comprehensive account of the linguistic diversities that may exist in this genre of writing. Following the same methodology as Biber's (1988), three new dimensions have been interpreted based on a factor solution matrix containing six factors. The data has been tagged for 180+ linguistic features, out of which, 121 have been shortlisted as significantly relevant to this research. A total number of 67 features contribute to the new factor solution. The new dimensions have been interpreted as **Abstract Information vs. Concrete Description**, **Contextualized Interactive Elaboration vs. Incoherent Informational Discourse**, and **Author's Personal Stance and Synthesis**, respectively. These new dimensions hold significance for future researchers who aim to highlight functional variation in their set of data as per these dimensions specified to argumentative essay writing. Elements of argumentation, specifically stance taking, may also be analyzed in argumentative texts by using this new model.

## REFERENCES

- Abdulaziz, M., Mahmood, M. A., & Azher, M. (2016). Variation in learner's argumentative essays - A multi-dimensional comparative analysis. *Science International*, 413-424.
- Asención-Delaney, Y., & Collentine, J. (2011). A multidimensional analysis of a written L2 Spanish corpus. *Applied linguistics*, 32(3), 299-322.
- Azher, M., & Mehmood, M. A. (2016). Exploring Variation across Pakistani Academic Writing: A Multidimensional Analysis. *NUML Journal of Critical Inquiry*, 14(2), 86.
- Azher, M., & Mehmood, M. A. (2016). Exploring Variation across Pakistani Academic Writing: A Multidimensional Analysis. *NUML Journal of Critical Inquiry*, 14(2), 86.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.
- Biber, D. [LU eLearning]. (2015, Oct 27). *Doug Biber: Multidimensional Factor Analysis lecture at UCREL Summer School 2015* [Video File]. Retrieved from <https://www.youtube.com/watch?v=TuFfRpnZ8s0>
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. 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. (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. doi:10.5054/tq.2011.244483
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (2007). *Longman Grammar of Spoken and Written English*. London: Longman.
- Crossley, S. A., Allen, L. K., & McNamara, D. S. (2014). A Multi-Dimensional analysis of essay writing – What linguistic features tell us about situational parameters and the effects of language functions on judgments of quality. In T. B. Sardinha & M. V. Pinto (Ed.), *Multi-Dimensional Analysis, 25 years on. A tribute to Douglas Biber* (pp. 197-237). Amsterdam: John Benjamins Publishing Company.
- Egbert, J. (2012). Style in nineteenth century fiction: A multi-dimensional analysis. *Scientific Study of Literature*, 2(2), 167-198.
- Egbert, J. (2014). Student perceptions of stylistic variation in introductory university textbooks. *Linguistics and Education*, 25, 64-77.
- Egbert, J. (2014). Student perceptions of stylistic variation in introductory university textbooks. *Linguistics and Education*, 25, 64-77.
- Egbert, J. (2015). Publication type and discipline variation in published academic writing - Investigating statistical interaction in corpus data. *International Journal of Corpus Linguistics*, 1-29.
- Egbert, J. (2015). Publication type and discipline variation in published academic writing - Investigating statistical interaction in corpus data. *International Journal of Corpus Linguistics*, 1-29.

- Friginal, E., & Weigle, S. (2014). Exploring multiple profiles of L2 writing using multi-dimensional analysis. *Journal of Second Language Writing, 26*, 80-95.
- Gardezi, S. A., & Nesi, H. (2009). Variation in the writing of economics students in Britain and Pakistan: the case of conjunctive ties. *Academic writing: at the interface of corpus and discourse, 236-250*.
- Gorsuch, R.L. (1974). *Factor analysis*. W.B. Saunders Company.
- Granger, S. (2002). A bird's-eye view of learner corpus research. *Computer learner corpora, second language acquisition and foreign language teaching, 6*, 3-33.
- Hymes, D. (1984). Sociolinguistics: stability and consolidation. *International Journal of the Sociology of Language, 1984(45)*, 39-46.
- Imtiaz, Z., & Mahmood, A. (2014). Genre Analysis of Argumentative Essays of Pakistani EFL Learners. *Journal of Education and Practice, 95-100*.
- Kachru, B. B. (1984). The alchemy of English: Social and Functional Power of Non-Native Varieties. *Language and Power, 176-193*.
- Kachru, B. B. (1985). Institutionalized Second Language Varieties. *The English Language Today, 211-226*.
- Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing, 20(2)*, 148-161.
- Mahboob, A. (2009). English as an Islamic language: A case study of Pakistani English. *World Englishes, 28(2)*, 175-189.
- Sajid, M., & Siddiqui, J. A. (2015). Lack of Academic Writing Skills in English Language at Higher Education Level in Pakistan: Causes, Effects and Remedies. *International journal of language and linguistics, 2(4)*, 174.
- Sarfraz, S. (2011). Error analysis of the written English essays of Pakistani undergraduate students: A case study. *Asian Transactions on basic & Applied sciences, 1(3)*, 29-51.
- Shaw, P. (2009). Linking adverbials in student and professional writing in literary studies: What makes writing mature. *Academic writing: At the interface of corpus and discourse, 215-235*.
- 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.
- Ure, J. (1982). Introduction: approaches to the study of register range. *International Journal of the Sociology of Language, 1982(35)*, 5-24.
- Van Rooy, B. (2008). A multidimensional analysis of student writing in Black South African English. *English World-Wide, 29(3)*, 268-305.
- Van Rooy, B. (2008). A multidimensional analysis of student writing in Black South African English. *English World-Wide, 29(3)*, 268-305.
- Xiao, R. (2009). Multidimensional analysis and the study of world Englishes. *World Englishes, 28(4)*, 421-450.