The ‘English’ Problem: MD Analysis of Reporting World in South Asian Newspapers

Sadia Ali (Main author)
Department of English, Air University E-9, Islamabad, 44000 Pakistan
diaali2010@gmail.com

Muhammad Ali (Co-author)
English Language Centre, Umm Al-Qura University, Makkah, Kingdom of Saudi Arabia
maattar@uqu.edu.sa

Muhammad Usman Ghani (Corresponding author)
Department of English, College of Science and Humanities, Prince Sattam Bin Abdulaziz University, Al Kharj, Kingdom of Saudi Arabia
m.ghani@psau.edu.sa

Biodata:
Sadia Ali is a Ph.D. in English Linguistics from Air University, Islamabad, Pakistan. She presented several papers at national and international conferences. Her research interests include ELT, Corpus linguistics, and World Englishes.

Muhammad Ali is a Ph.D. in English Linguistics. Besides his several years of work at Pakistani universities, he has been teaching at Umm-Al Qura University, Makkah, Saudi Arabia since 2015. Apart from his teaching and research experience, he has developed materials and designed syllabus for learners and teachers of English at International Islamic University. His work has been published in reputed national and international journals. His major publications include areas of World Englishes, pure and applied linguistics, ELT, corpus linguistics, and linguistic variations.
Muhammad Usman Ghani is working as an Assistant Professor at the Department of English, College of Science and Humanities, Prince Sattam Bin Abdulaziz Univ Al-Kharj Saudi Arabia. His interests includes South Asian Literature and sociolinguistic problems and patterns in South Asia and the Middle East.

Abstract
This study seeks to analyze linguistic variations among the varieties of English used in press in South Asia, i.e., Pakistan, India and Bangladesh (erstwhile pre-partition British India) in comparison with British English. We argue that these three distinct varieties of English have effectively evolved in this varied region where English was introduced through colonization. A corpus of World Reportage (WR) of 288 texts was analyzed using Biber’s (1988) multi-dimensional model. Regression method was used for the computation of factor scores while analysis of variance (ANOVA) was performed to find the linguistic variations in the texts. The results provide substantial evidence that each country has developed its style of WR. Notably, however, Indian WR falls closer to the British WR to produce informational, narrative, non-argumentative and abstract discourse whereas Bangladeshi WR falls closer to British WR to produce explicit discourse. Further, Pakistani WR is found to be a distinct variety in comparison to all. The findings suggest that English language teaching, particularly ESP textbooks should include various varieties of English so that the journalism students are exposed to various writing styles of WR. The corpus developed for this study can also serve as the basis for designing curriculum and textbooks for journalism students.

Keywords: ESP, Multi-dimensional analysis, press reportage, register analysis, South Asian English, world reportage

Introduction
Pakistan, Bangladesh and India share interesting history: Pakistan and India appeared as independent countries after winning independence from the British Raj in 1947 and Bangladesh got independence from Pakistan in 1971. As these countries were one before separation, they share largely the same cultural, ethnic, religious, linguistic and geographical backgrounds (Hickey, 2009). They also share the early history of the usage of English and its origin with one another. English was first introduced to these South Asian countries through colonization. During the British Raj, English gained political domain and socio-economic status so well that even after independence both the countries India and Pakistan (including Bangladesh known
as East Pakistan then) maintained the role of English in the new states (Mahboob, 2003). English, which initially spread in British India because of socio-political and economic factors, continued to enjoy the status of one of the main languages in this part of the world. However, on later stages, various other factors like the influence of indigenous languages (Schilk et al., 2012), non-native English language teachers, religions, local cultures, etc., started indigenizing English in the sub-continent.

There are certain attempts at studying the localized Englishes: South Asian English (e.g., Tickell, 2016; Kachru et al., 2009; Brians, 2003) and to be more precise Indian English (Nandi, 2013; Gargesh, 2008, etc) and Pakistani English (Shakir and Deuber, 2018; Asghar et al., 2018) in comparison with British English. However, most of the research studies (like Islam, 2018; Quayum & Hassan, 2018; Huq, 2018) on Bangladeshi English are confined to English language teaching only. There are some attempts at comparing Pakistani and Indian Englishes (e.g., Taimur-ul-Hassan and Seyal, 2016; Saffee, 2016; Hussain, 2015; Batang, Dayag–Vecaldo, & Medriano Jr, 2018). In addition, previous studies have largely been limited to individual linguistic features to comparing one variety of English with British or American English. However, there is a need to explore linguistic variation which occurs in ‘one’ variety of English learnt at the same origin/ source when it crosses the borders of Pakistan, India and Bangladesh.

The present study, therefore, is an attempt to explore the linguistic variations in written corpora of the selected South Asian regions, i.e., Pakistan, Bangladesh and India, which remained British colony known as the Indian subcontinent/ British India. The study uses Biber’s multi-dimensional (1988) approach as a framework which investigates overall patterns of linguistic variations.

**A Review of the Studies on South Asian Englishes (SAE)**

English as an international language is used as a shared means of communication also known as lingua franca throughout the world. In most of the South Asian countries, the English language was introduced via colonization. However, with the passage of time, different varieties of English have emerged, roughly defined by the national borders between South Asian nations i.e., Bangladeshi English, Indian English, Nepali English, Pakistani English, Sri Lankan English. With the emergence of Englishes, researchers in South Asia started paying attention to the features associated with new Englishes. Many researchers worked on South Asian Englishes (e.g., Kirkpatrick, 2014; Schilk et al., 2012; Kachru & Smith, 2009; Baumgardner, 1996; Go Silk et al. 2020; Bedu, 2020)and based on their findings, these Englishes were considered distinct varieties.
According to Hickey (2009: 536) “[t]he dominance of English in the public lives of South Asia (except Nepal and Bhutan, which were not part of the empire) is a legacy of the British colonialism with its administration and the establishment of English in their educational system”. He further adds that South Asia is a ‘socio-linguistic’ area which shared ‘phonological, syntactic, lexical and stylistic features’ to a great extent. Some elements influence the English of a particular country. Kachru et al., (2009: 178) assert, “Although it is fairly homogeneous across the region, sharing linguistic features and tendencies at virtually all linguistic levels, there are also differences based on various factors”. Non-native English language teachers, religions and local cultures indigenizing English in the sub-continent are some of the factors that influence English language of a region. Kirkpatrick (2014) adds that the mother tongue of a speaker is also an important factor that influences his/her English. Estai and Savarabadi (2020) are of the view that EFL teachers with regard to EIL must be culturally aware. According to Schilk et al., (2012: 139), “[i]t is obvious that there are also historical and functional differences between South Asian Engishes contributing to the manifestation of linguistic variation across varieties of English in the sub-continent”. So, with the emergence of the new varieties of English, these indigenized Engishes became the centre of the attention of many researchers. British Empire in South Asia (now Pakistan, Bangladesh and India, erstwhile British colony) is one of them.

The works of Rahman (1990), Baumgardner (1993) and Tallat (2002) are considered to be pioneering works on the Pakistani variety of English, but they are mostly impressionistic and attempt to establish Pakistani English as a distinct variety on the basis of individual linguistic features. Biber (1988) introduces a more comprehensive approach to study the co-occurrence of linguistic features instead of individual linguistic aspects. Many researchers used this approach to study registers throughout the world. For instance, Thompson et al., (2017) conducted a multi-dimensional analysis to explore the linguistic features in the corpus of journal articles in environmental studies. Egbert (2015) investigated the linguistic variations across three publication types, i.e., academic books, university textbooks and journal articles from two disciplines, i.e., Biology and History using Biber’s multi-dimensional modal. There are a few studies that used multi-dimensional modal on the languages other than English. For example, Biber and Hared (1994) conducted multi-dimensional analysis on Somali language. Likewise, Kim and Biber (1994) conducted MD modal on Korean, Jang (1998) on Taiwanese, Biber et al., (2006) and Ascencon-Delaney (2014) on Spanish and Sardinha et al, (2014) on Portugese language. Moreover, Burt and Bauer (1996) and Westin and Geisler (2002) conducted a diachronic study of press editorials by using multi-dimensional modal.
With register coming into prominence as a major predictor of linguistic variation (Atkinson & Biber, 1994), various studies explored a number of registers to establish it as a separate/independent variety of English. Researchers in Pakistan (e.g., Shakir and Deuber, 2019; Asghar et al., 2018; Shakir and Deuber, 2018) also used the model of co-occurring linguistic features and worked on various registers of English used in Pakistan establishing it as an independent variety of English.

As far as Bangladesh is concerned, there is no significant study which focuses on Bangladeshi English as a separate variety. Most of the studies (e.g., Islam, 2018; Sultana, 2014; Rahman, 2005; Arifa, 2000) either investigated the attitude of students in learning English language or teaching English skills including the methods for English language teacher’s education.

Kachru (1997) identifies a number of Indian English features which are common in South Asian varieties of Englishes. The results of his study show that a sub-variety or register could be used according to its language function. Though some studies compare Pakistani and Indian English such as (Hassan and Seyal, 2016; Saffee, 2016; Sajjad, 2015; Uzair et al., 2012) yet no study analyzes the linguistic features of Pakistani, Indian and Bangladeshi Englishes in comparison with British English to find out whether after independence these countries maintained the role and the usage of English in the new states or developed their own varieties of Englishes. This study aims at analyzing the WR of Pakistani, Indian and Bangladesh press in comparison with British WR. In other words, it attempts to analyze these Englishes in comparison with British English in order to explore how far the varieties of English used in Pakistan, India and Bangladesh are similar to or different from British English.

**Corpus Compilation and Methods**

The world news category within press reportage was selected as the register in focus of this study. The corpus developed for this study consisted of 288 texts (332,791 words). Data was collected from four leading newspapers from Pakistan, India and Bangladesh (erstwhile British colony) and Britain. 72 texts from each county were taken as a sample for the study (see table 1 for detail). The length of the texts varies from 800 to 1200. Their frequencies were normalized to per 1,000 words.
Table 1. Text sample and word counts in World Reportage corpus

<table>
<thead>
<tr>
<th>Country</th>
<th>Text &amp; Word counts</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pakistan</em></td>
<td>72 (88,288)</td>
</tr>
<tr>
<td><em>India</em></td>
<td>72 (77,734)</td>
</tr>
<tr>
<td><em>Bangladesh</em></td>
<td>72 (77,529)</td>
</tr>
<tr>
<td><em>Britain</em></td>
<td>72 (89,240)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>288 (332,791)</td>
</tr>
</tbody>
</table>

*Word counts are in parenthesis

Biber’s (1988) multi-dimensional model (for details see appendix 1) was used as a framework for this study. Biber (1988) tagger was used to grammatically annotate the World Reportage corpus. In order to assign grammatical categories to the lexical items, all the text files (288) were run through the Biber tagger. Biber (1988) identified 67 linguistic features. These 67 linguistic features were categorized in 16 grammatical categories, viz, (A) tense and aspect markers, (B) place and time adverbial, (C) pronouns and pro-verbs, (D) questions, (E) nominal forms, (F) passives, (G) stative forms, (H) subordination features, (I) prepositional phrases, adjectives, and adverbs, (J) lexical specificity, (K) lexical classes, (L) modals, (M) specialized verb classes, (N) reduced forms and dispreferred structures, (O) coordination, and (P) negation (Biber, 1988: 72). (For detail, see appendix 2)

For the computation of factor scores, the regression method was used in this study. The last step within the quantitative analysis was to quantify the linguistic variations among the selected countries. For this purpose, analysis of variance (ANOVA) was performed. Finally, the communicative functions associated with each dimension were interpreted based on the linguistic composition of the dimension.

**Linguistic Analysis of World Reportage**

This section presents a quantitative and functional interpretation of World Reportage (WR) of Pakistani, Indian and Bangladeshi English press on Biber’s (1988) five textual dimensions. It also compares WR of these countries and that of Britain in the linguistic space which co-occurring patterns define providing deep insight to the teachers of writing in journalism and media studies. So far as dimension 1 is concerned, it is labeled as Involved vs. Informational discourse. The linguistic features on positive polarity perform the function of involvedness. On the other hand, linguistic features on negative polarity together produce informational discourse. Different factors are linear combinations of the original variables and they represent different groupings of high frequency co-occurring linguistic features. Correlation can be
positive or negative indicating the degree to which two linguistic features vary in relation to each other. While a large positive correlation shows that the linguistic features systematically occur together, a large negative correlation indicates a complementary distribution of linguistic features specifying that the linguistic features covary, in a manner that if one linguistic feature is present, the other must be absent (Biber, 1988).

Figure 1 shows that all the countries produce informational discourse in their WR, however, there is significant statistical difference among the mean values of all the countries. Linguistic features like nouns, preposition and attributive adjective produce informational discourse. In comparison with WR of other countries, Bangladeshi WR, with mean score (-19.82), produces the highest informational discourse. Pakistani WR with mean score (-19.06), is slightly less informational than Bangladeshi WR. Among these three countries, Indian WR register produces the least informational discourse. Figure 1 shows that the mean score of British WR is -14.9. When WR category of these South Asian countries is compared to that of Britain, Indian WR, with mean score (-17.14), has been found the closest to that of Britain in producing informational discourse on dimension 1. The difference between British and Bangladeshi WR register is the highest.

![Figure 1. Linguistic variation of World Reportage of Britain, Pakistan, India and Bangladesh on dimension 1](image)

Nouns, prepositions and attributive adjective are all informational in one way or the other (Biber, 1988). The normalized mean score of *nouns* in Pakistani, Indian and British WR (351.91, 352.68 and 345.35 respectively) is quite low as compared to Bangladeshi WR (360.52). Further, attributive adjectives elaborate nominal information and pack information using relatively less words and structures (Sardinha & Pinto, 2014). In Bangladeshi WR, attributive adjectives with normalized mean frequency of (57.70) show the highest score in comparison with Pakistani (56.52), Indian (53.77) and British WR (50.19) suggesting that
Bangladeshi WR is the most informational among the selected countries. Preposition is one among the other linguistic features that produce informational discourse. Preposition is considered to be “a device for integrating information into idea unit” (Biber et al., 2007: 74) and prepositional phrases serve to integrate high amounts of information into a text (Biber, 1988). The normalized mean scores of prepositions in Pakistani WR (126.23) is close to Bangladeshi WR (125.58). The mean normalized frequency of prepositions in Indian WR (120.95) and British WR (119.53) are considerably less than Pakistani WR and Bangladeshi WR.

In example (1), taken from Indian newspaper, *The Hindu*, the italicized words are the examples of linguistic features that together perform a function of producing information in the text. In the following example, nouns like *research, equipment, ICBM, rocket, Kim, year, day, country, UN, missile, sanctions, month, Pyongyang, test and Weapons* produce informational discourse. In the likewise manner, preposition like *in, under and on* also produce informational discourse. Moreover, the presence of attributive adjectives like *cutting edge, actively progressing, last stage, New year’s Day, last month, largest nuclear test, successful ICBM test, significant step and secretive Pyongyang’s weapons* enhance the presence of informational discourse.

(1) *Research and development of cutting edge arms equipment is actively progressing and ICBM rocket test launch preparation is in its last stage,” Mr. Kim said during a televised New Year’s Day speech. The country has been under UN sanctions since 2006 over its nuclear and ballistic missile tests. The sanctions were tightened last month after Pyongyang conducted its fifth and largest nuclear test on Sept. 9. A successful ICBM test launch would mark a significant step forward for secretive Pyongyang’s weapons capability. (The Hindu, India, January 1, 2017.)*

Linguistic features like nouns (*China, Neighbours, Brunei, Malaysia, Philippines, Taiwan, Vietnam, Xi, head, serious, complex, Zhang Zhijun*), preposition (*in*) and attributive adjectives (*South China Sea, self-ruled Taiwan, New Year’s greetings, New Year’s message, uncertain factors, Taiwan Strait*) together perform the function of producing Informational discourse. Example (2) is taken from British newspaper, *Daily Mail*. It includes informational linguistic features. The italicized words produce Informational discourse of British press reportage.

(2) *China claims most of the South China Sea. Neighbours Brunei, Malaysia, the Philippines, Taiwan and Vietnam also have claims. While Xi made no direct mention*
of self-rulled Taiwan, aside from extending New Year's greetings to them, the head of China's policy-making Taiwan Affairs Office in his New Year's message said that 2017 would see uncertainty. "Looking ahead to 2017, the situation in the Taiwan Strait is complex and serious, and the development of relations are facing many uncertain factors and risk," Zhang Zhijun said. (Daily Mail, Britain, January 1, 2017)

Dimension 2 is labelled as Narrative vs. Non-narrative concerns. While, positive linguistic features including past tense verb, third-person pronoun, verb-perfect aspect and public verbs produce narrative discourse, negative linguistic features produce non-narrative discourse. On this dimension, all the countries show narrative concern in their world press reportage. Figure 2 shows that there is a significant statistical difference between the mean scores of all the countries. In comparison with WR of other countries, WR of Bangladesh (2.05) is the most narrative. Pakistan (1.8) shows slightly less narrative discourse production than Bangladesh. The figure shows that Britain with mean score of (0.8) is the least narrative as compared to the selected South Asian countries. When the mean scores of WR of these South Asian countries are compared to that of Britain, India with the mean score of (1.55) is found closest to Britain in producing narrative discourse. In contrast, the difference between British and Bangladeshi WR is the highest.

Figure 2. Linguistic variation of World Reportage of Britain, Pakistan, India and Bangladesh on dimension 2

Past tense verbs, public verbs, and third-person pronoun can be considered as primary markers of narrative discourse (Biber, 1988). Past tense verbs describe past events. The mean normalized scores of past tense verb in Pakistani, Indian and British WR (47.15, 43.48, 41.7, respectively) are less than the scores of Bangladeshi WR (48.50). Bangladeshi WR, therefore,
is the most narrative amongst Pakistani, Indian and British WR. In addition, to past tense verbs, public verbs are also the markers of indirect speech (Biber, 1988). Bangladeshi WR with a high mean normalized score of public verbs (15.97) is more indirect in comparison with Indian (13.21) and British (10.71) WR. However, it is relatively close to Pakistani WR (15.24). Further, third-person pronouns mark reference to animate, especially human (Orsan & Evans, 2001) and narrative discourse heavily depend on it. In British WR, the mean normalized frequency of third-person pronoun (28.20) is the highest in comparison with Indian (24.56), Pakistani (18.57) and Bangladeshi (18.30) WR.

In example (3), the italicized words are the linguistic features that produce narrative discourse. Past tense verb is the main marker of narration. They narrate the events that took place in the past. Past tense verbs like left, expressed, continued, was responsible, said and pressed on produce narrative discourse. The presence of third-person pronoun like she, he, his, and public verbs like expressed, remarks, said also make the text narrative.

(3) U.S. President-elect Donald Trump on Saturday left open the possibility of meeting with Taiwan’s President if she visits the United States after he is sworn in on Jan. 20 and also expressed continued scepticism over whether Russia was responsible for computer hacks of Democratic Party officials. In remarks to reporters upon entering a New Year’s Eve celebration at his Mar-a-Lago estate, Mr. Trump said, “We’ll see,” when pressed on whether he would meet Tsai Ing-wen, Taiwan’s President if she were to be in the United States at any point after he becomes President. (The Hindu, India, January 1, 2017)

In example (4) the linguistic features like past tense verb (disagreed, sought, born, allowed, conducted, gathered), third-person pronoun (who) and public verbs (suggested, clarify) show the dense presence of linguistic features that produce narrative discourse. Moreover, verb-perfect aspect marks action in past time with current relevance. In example (4), verb-perfect aspect like has suggested produces narrative discourse.

(4) Most Americans, however, don't think the country should show a preference for Christian refugees, as Trump has suggested. Some 56 percent, including 72 percent of Democrats and 45 percent of Republicans, disagreed that the country should "welcome Christian refugees, but not Muslim ones." On Tuesday, the Trump administration sought to clarify that citizens of U.S. ally Israel who were born in Arab
countries would be *allowed* into the United States. The Reuters/Ipsos poll was *conducted* online in English in all 50 states. It *gathered* poll responses from 1,201 people including 453 Democrats and 478 Republicans. (Daily Mail, Britain, February 1, 2017)

So far as dimension 3 is concerned, it is labelled as Explicit vs. Situation dependent discourse. The linguistic features on positive polarity like wh-pronoun-relative clause-object position, wh-pronoun-relative clause-subject position, pied-piping, coordinating conjunction and singular noun-nominalization produce explicit discourse. Whereas, negative linguistic features like adverb of time and adverb of place produce situation-dependent discourse. Interestingly, all linguistic features show explicit discourse production in WR. Figure 3 shows that there is a significant statistical difference among the mean values of all the countries. In comparison with WR of the other countries, Indian WR, with the mean score of (4.95), is the most explicit. Pakistan (4.7) is slightly less explicit than India and Britain (3.26) is the least explicit among the selected countries. When WR category of these South Asian countries is compared to that of Britain, Bangladesh (3.54), in terms of its mean score, is found closest to Britain. In contrast, the difference between British and Indian WR is the highest.

![Figure 3. Linguistic variation of World Reportage of Britain, Pakistan, India and Bangladesh on dimension 3](image)

Noun-nominalization and pied-piping are considered to be devices for explicit identification of referents in a text (Biber, 1988). Shireen (2017) finds that singular noun-nominalization is used to integrate information into fewer words. Nominalization, which is a primary marker of producing explicit discourse, occurs more frequently in Indian WR (68.29) in comparison with Bangladeshi (66.45), Pakistani (65.04) and British (54.92) WR. Further, pied-piping constructions are found more in written discourse than spoken discourse (Biber et al., 2007).
Each country shows the significant mean normalized frequency of pied-piping in their WR. Interestingly, Pakistani, Bangladeshi and British WR have the same normalized frequency of pied-piping (i.e., 0.49). However, the mean normalized frequency of pied-piping in Indian WR (0.38) is considerably less than Pakistani, Bangladeshi and British WR.

Example (5), from Bangladeshi newspaper, Daily Observer, shows the presence of noun-nominalization like realising, developing, threaten and Wh clauses like Kim Jong-Il who died and Donald Trump, who took to Twitter, produce explicit discourse.

(5) Analysts are divided over how close Pyongyang is to realising its full nuclear ambitions, but all agree it has made enormous strides since Kim took over as leader from his father Kim Jong-Il who died in December 2011. Kim Jong-Un said in a New Year’s speech that Pyongyang was in the “final stages” of developing an intercontinental ballistic missile of the kind that could threaten US territory. The address drew a swift response from US president-elect Donald Trump, who took to Twitter vowing to halt Pyongyang in its tracks. (Daily Observer, Bangladesh, January 2, 2017)

Excerpt (6) from Daily Mail exhibits a high frequency of nominalization (security, operation, fugitive assailant, foreigners, reporters, coalition, militants and borders) and relative clauses (who gave her name, who was in the club and all three of us heard that).

(6) "At first we thought some men were fighting with each other," said a Lebanese woman who gave her name as Hadeel and who was in the club with her husband and a friend…“We heard the guy screaming Allahu Akbar (God is Greatest), all three of us heard that ... A massive security operation unfolded to track down the fugitive assailant or assailants and any conspirators…Reina were foreigners but only 21 bodies had so far been identified. He told reporters 69 people were in hospital, four of them in critical condition. Turkey is part of the U.S.-led coalition against Islamic State and launched an incursion into Syria in August to drive the radical Sunni militants from its borders. (Daily Mail, Britain, January 1, 2017)
On dimension 4, linguistic features like *infinitive verb, modal of prediction, suasive verb, subordinating conjunction-conditional, modal of necessity* and *adverb within auxiliary* on positive polarity produce Argumentative discourse. This dimension is labelled as Overt Expression of Argumentation/Persuasion. This dimension does not have any negative linguistic feature in Biber’s (1988) study. However, in the present study *pronoun it, private verbs* and *hedges* produce non-argumentative discourse. Figure 4 shows that all the selected countries produce non-argumentative discourse in their WR, however, there is a significant difference between the mean scores of all the countries. In comparison with WR of other countries, Pakistan (-2.62) is the most non-argumentative. India with a mean score of (-2.37) is less non-argumentative than Pakistan. Bangladesh (-2.33) is the least non-argumentative among the other countries. Like dimension 1 and 3, on this dimension also India, in its mean score, is the closest to Britain (-2.47). However, the difference between British and Pakistani WR is the highest on this dimension.

Pronoun *it* marks inexplicitness and a non-informational focus in a discourse (Biber et al., 2007). The mean frequency of pronoun *it* in Pakistani WR (8.75) is quite high in comparison to its frequency in Bangladeshi (8.54), Indian (8.53) and British (7.95) WR, which suggests that Pakistani WR is the most non-argumentative among them. Biber (1988: 105) asserts, “Private verbs are used for the overt expression of private attitudes, thoughts, and emotions”. They are the main markers of producing non-argumentative discourse. The mean normalized scores of private verbs in Pakistani (5.83), Bangladeshi (6.05), Indian (6.35) and British WR (8.06) suggest that these countries produce non-argumentative discourse in their WR. Hedges
are markers of uncertainty (Biber, 1988). The mean normalized frequency of hedges in British WR (0.35) is considerably high in comparison with Pakistani (0.24), Bangladeshi (0.19) and Indian (0.27) WR. Example (7) has been taken from the Indian newspaper, The Telegraph. It shows the dense presence of pronoun it and private verbs like warned and hedges like but. It indicates that WR from the Pakistani newspaper shows the presence of linguistic features which produce non-argumentative discourse.

(7) Actress Meryl Streep used her acceptance speech for the Cecil B. DeMille Award at the Golden Globes to call out President-elect Donald Trump for mocking a disabled New York Times reporter in 2015, and warned that a free press would need to be defended. But there was one performance this year that stunned me. It sank its hooks in my heart. Not because it was good; there was nothing good about it. But it was effective and it did its job. It made its intended audience laugh, and show their teeth. It was that moment when the person asking to sit in the most respected seat in our country imitated a disabled reporter. (The Telegraph, India, January 1, 2017)

Example (8) is taken from British newspaper, Daily Mail. In this example the words like it, urged and address are the features of non-argumentative discourse.

(8) The statement says anti-Semitic attitudes and national security fears were among the reasons for turning away Jewish refugees. The museum says there are legitimate refugees fleeing Syrian President Bashar Assad's regime and genocidal acts by the Islamic State group...Its statement urged that U.S. policy should address security concerns while protecting legitimate refugees. (Daily Mail, Britain, February 1, 2017)

Dimension 5 is labelled as Impersonal (abstract) vs. Non-impersonal (Non-abstract) style. The linguistic features on positive polarity like adverb-conjunct, agentless passive verb, passive verb, passive postnominal modifier and subordinating conjunction show the abstract or impersonal style in text. Linguistic features type/token ratio on negative polarity show the non-abstract or non-impersonal style of discourse production. Figure 5 shows that all the countries in their WR show abstract style of writing, however, there is a significant statistical difference between the mean values of all the countries. In comparison with WR of other countries, Bangladeshi WR (1.79) is the most impersonal. India with a mean score of (1.33) is less
impersonal than Bangladesh. Pakistan (1.28) is the least abstract among the other selected South Asian countries.

Figure 5. Linguistic variation of World Reportage of Britain, Pakistan, India and Bangladesh on dimension 5

Passive constructions, conjuncts and agentless passive verbs present abstract information (Biber, 1988). Passive constructions dropped agent, thus resulting in a more abstract discourse. In Bangladeshi WR, passives with a normalized mean frequency of (1.95) show the highest score when compared to Pakistani, Indian and Britain WR (1.71, 1.77 and 1.67, respectively). This comparison indicates that WR of Bangladesh is the most abstract among the selected countries. Conjuncts mark the logical relationship between clauses (Shaves, 2012) and they are the main marker of producing abstract discourse. The mean frequencies of adverbial conjunct in Pakistani (1.76), Bangladeshi (1.80), Indian (2.19) and British WR (2.16) indicate that the WR of these countries produces abstract discourse. “Discourse with very frequent passive constructions is typically abstract and technical in content and formal in style” (Biber, 1988: 112). In British WR, passive verbs with a normalized mean frequency of (9.69) show the highest score in comparison with Pakistani (8.68), Bangladeshi (8.93), Indian (8.43) WR.

When WR category of these South Asian countries is compared to that of Britain, India is found the closest to Britain (1.53) in producing impersonal discourse on dimension 5. The difference between British and Bangladeshi WR register is the highest. See extract (9) for example:

(9) Antonio Guterres took the reins of the United Nations on New Year’s Day, promising to be a “bridge-builder” but facing an antagonistic incoming U.S. administration led by Donald Trump who thinks the world body’s 193 member states do nothing except talk and have a good time. Mr. Guterres said the values enshrined in the U.N. Charter that
should define the world that today’s children inherit peace, justice, respect, human rights, tolerance and solidarity are threatened, “most often by fear.” (The Hindu, India, January 2, 2017)

In example (9), the italicized words highlight the linguistic features like *administration led by Donald Trump, most often by fear* that produce abstract style of information.

(10) I understand that the *country was founded on immigrants,*" said Hoffman, who participated in the poll. "Please, I get that. But I’m worried that refugees are coming in and *being supported by my tax dollars.*"…Most Americans, however, don't think the country should show a preference for Christian refugees, as Trump has suggested. Some 56 percent, including 72 percent of Democrats and 45 percent of Republicans, disagreed that the country should "welcome Christian refugees, but not Muslim ones. (Daily Mail, Britain, February 1, 2017)

In example (10), taken from *Daily Mail,* the presence of words like *country was founded, being supported by my tax dollars and however* are the examples of linguistic features associated with abstract style of discourse.

**Conclusion & Future Direction**

As referred to in the beginning, the current study compared the WR of the three countries, i.e., Pakistan, India and Bangladesh (where English was introduced through colonization) with British WR on Biber’s (1988) five textual dimensions. On dimension 1, all the countries produce Informational discourse, however, there is a marked difference between the mean scores of Bangladeshi and British WR. As far as British and Indian WR are concerned, their production of Informational discourse is quite similar. On dimension 2, although all the countries produce narrative discourse, the difference between Bangladeshi WR and British WR is the highest. Indian WR falls quite close to British WR in its mean score. So far as dimension 3 is concerned, all the countries produce explicit discourse in their WR. However, on this dimension, Indian WR demonstrates a great variance in its mean scores from British WR. British WR and Bangladeshi WR are almost similar in producing explicit discourse. On dimension 4, all the countries produce non-argumentative discourse in their WR. On this dimension also, there is less difference between the mean scores of British and Indian WR and there is a large difference between the mean scores of British and Pakistani WR. On dimension
5, all the countries show abstract style in their WR. Again British WR and Indian WR are quite common in their abstract style of WR. However, there is a significant difference between the mean scores of British and Bangladeshi WR.

Overall, the results of World press reportage show that on D1, D2, D4 and D5 Indian WR and British WR are close in producing informational, narrative, non-argumentative and abstract discourse. Only on D3 Bangladeshi WR is close to British WR in producing explicit discourse. Pakistani WR is significantly different from British WR on all the five dimensions. Although English language was introduced to all the three countries, Pakistan, India and Bangladesh, through British colonization, the results of the study provide substantial evidence that each country has developed its style of newspaper reportage, particularly WR. However, it is important to note that Indian WR is quite close to British WR. Further, Pakistani WR is a distinct variety as compared to Indian, Bangladeshi and British WR.

Future researchers may observe linguistic variations among the other sub-categories of press reportage of the selected countries or other South Asian countries. Further, this study is corpus-based. It might also be possible to go beyond the findings of the present study and probe the reasons behind the identified linguistic variation and dimensions. The findings and corpus of this study are useful for syllabus designers of English for specific purpose. While ESP teachers are well aware of the fact that the register of journalism is conspicuously different from other registers, they also need to be aware that this register shows further variation as each country has developed its own norms which is demonstrated through the findings of this study. Initially, the learners are required to follow the traditional models of journalistic writing. However, on later stages, they must be exposed to the varieties of journalistic writings developed in different countries through corpus-based English language teaching so that they can analyse the actual patterns of writing in journalism.

Acknowledgment

The authors would like to thank the Deanship of Scientific Research at Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia, for the support in the publication of this manuscript. We are also thankful for the assistance provided by Biber's lab at North Arizona University, USA for tagging the data.
References


Appendix 1
Co-occurring Linguistic Features on Five Textual Dimensions of 1988 MD analysis of Press Reportage

**Dimension 1: Involved vs. Informational Discourse**

**Positive Feature**
- Private Verbs
- ‘That’ deletion
- Verb (uninflected present, imperative & third Person)
- Second Person pronoun/Possessive
- Verb ‘Do’
- Demonstrative Pronoun
- Adverb/Qualifier-Emphatic (e.g., just, really)
- First person pronoun/possessive
- Pronoun ‘it’
- Verb ‘Be’ (uninflected present tense, verb, and auxiliary)
- Subordinating Conjunction-Causative
- Discourse Particle
- Nominal Pronoun
- Adverbial –Hedge
- Adverbial/Qualifier-Amplifier
- Wh-question
- Modals of Possibility
- Co-ordinating conjunction-clausal
- connector
- Wh-clause
- Stranded Preposition

**Negative Features**

**Dimension 3: Explicit Vs. Situation Dependent Discourse**

**Positive Feature**
- Wh-pronoun-relative clause-object position
- Wh-pronoun-relative clause-subject-position
- Wh-pronoun-relative clause-object position
- With prepositional fronting (pied-piping)
- Nominalization
- Coordinating Conjunction –phrasal connector
- Singular noun-nominalization

**Negative Features**
- Adverb of time
- Adverb of Place
- Adverb Other

**Dimension 4: Overt Expression of Argumentation /Persuasion**

**Positive Features**
- Infinitive Verb
- Modal of Prediction
- Persuasive Verb
- Subordinating conjunction-conditional
- Modal of Necessity
- Adverb within auxiliary

**Negative Features**
- (no negative features)
Nouns (excluding gerund)
Preposition
Attributive Adjective

**Dimension 2: Narrative vs. Non narrative**

**Concerns**

**Positive Feature**
Past Tense Verb
Third person pronoun (except ‘it’)
Verb-perfect Aspect
Public Verbs

**Negative Features**
(No negative Features)

---

**Dimension 5: Impersonal (Abstract) VS. Non impersonal (Non-Abstract Style)**

**Positive Features**
Adverbial-conjuncts
Agentless Passive verb
Passive verb + by
Passive Post nominal modifier
Subordinating conjunction-Other

**Negative Features**
(no negative features)