

Newspaper Racial Agenda: A Critical Discourse Analysis

Omama Tanvir¹

MPhil Scholar, Air University, Islamabad

Khoula²

MPhil Scholar, Air University, Islamabad

Syeda Tehseen Zahra³

Assistant Professor, Air University

Abstract

This study aims at researching the role of print media in creating a racial project through the development and perpetuation of ethnic stereotypes. For this purpose, American newspaper coverage of mass shootings occurring from year 2000 onwards are analysed, through corpus-based approach. The research focuses on a comparison between the discriminated reporting of white American shooters and the American shooters of different ethnic backgrounds by a close analysis of the lexical items, particularly adjectives. Thus, two corpora of 38,582 words are formulated based on sixteen mass shooting incidents. Among them eight incidents were carried out by white shooters and eight by non-white American shooters. The news articles are taken from the top American newspaper agencies such as *The Guardian U.S Edition*, *The New York Times*, *New York Daily* and *The Washington Post*. Critical Discourse Analysis approach presented by Fairclough (1992) and the concept of “Other” put forth by Edward Said (1978) serve as a theoretical framework for the study whereas Corpus Linguistics along with CDA serve as the two methodological constructs. The data is analyzed using software tools such as Wordsmith Tools and LancsBox 32. The findings reveal that the ethnicity of a shooter such as Korean, Afghan, Bosnian, Saudi and Vietnamese determines the shape a discourse would take. However there is a deliberate omission of the race when the perpetrator is a white man. The findings further reveal the lexical items with negative connotations like “terrorist”, “extremist” and “radical” are mostly used with Muslim perpetrators or with perpetrators of immigrant status to some extent but never with white mass murderers. These white killers are justified or humanized by declaring mental health issues as the root of the problem. The analysis yields that the articles about the white American perpetrators are mostly constituted of objective facts i.e. about the objective details of the incidents and the lives of the shooters. On the other hand, the articles about shooters with different ethnic background focus on the subjective details about the lives of the killers. Similarly adjectives with negative connotations such as ‘crazy, mad and abusive’ are never associated with white American shooters.

Keywords: adjectives, CDA, corpus linguistics, corpus-based, identity construction, racial conflicts, stereotypes, wordlist

1. Introduction

The past couple of decades have seen a great upsurge of crimes and terrorism all over the world. These acts of terrorism are manifested in different forms of violence. Most recent of this manifestation is mass shooting. Mass shootings are differentiated from other forms of violence by the fact that multiple victims are chosen apparently at random by the perpetrator, and where guns are the principle weapons used in the attack. Mass shootings have increased in United States of America in the past decade (Frisby, 2017). Americans and other people around the world who do not have firsthand experience with such incidents, come to know about these events through news media; especially television and newspapers (DeFoster, 2015). However, media cannot be taken as reliable and innocent tool of representation, production and consumption (Dijk, 1995). Authorities and elites use it as a tool in order to control the masses, and influence their attitudes (Dijk, 1995). It also helps in propagating “Us vs Them” dichotomy which imitates the attitude of white people towards other minority groups (Chuang, 2012). The dominant agenda propagated by the media is then received and absorbed by the masses as common sense, intensifying the disparity between white American people and other minority groups (Frisby, 2017).

The purpose of this study, however, is to investigate the possibility that American news coverage of mass shootings might constitute a racial projection with reference to the juxtaposition of coverage of American White perpetrators and perpetrators of different racial background. The prominence given to the race of the shooter when he is a Non-White American and the exclusion of race when he is a White member of the American community in the articles of top American news agencies may signal a racial project. This reinforces the foreigner status of the non-*whites*. The stereotypes reiterated through these articles contribute in strengthening the opposing *outsider or other* image, resulting in biased treatment of people with different ethnic background.

2. Literature Review

The aim of this study is to critically analyze the prominence given to the race of the shooter when he is a non-white American and the omission of race when he is White member of the American community, within the news articles. For this purpose works of different theorists in the broad interdisciplinary field of *CDA* have been elaborated. “Originally the word discourse is coined from a Latin word *discursus* which means conversation/stretch of talk” (Nawaz, Bilal, Kalsoom, Fayyaz & Nayyar, 2013, p. 330). Discourse is a piece of conversation whether formal or informal, but when it comes from a powerful source it takes on various other dimensions. In a political context, discourse plays a pivotal role to garner public support for many implicit or explicit agendas. Discourse is thus an exploitative tool which requires due attention and critical analysis from linguists. Otherwise the discursive practices would remain concealed.

To delve beneath the surface of discourse certain critical tactics and theories have emerged in the field of linguistics, critical discourse analysis being one of them. One of the preliminary proponents of the said theory is Dijk (1998) who defines Critical

Discourse Analysis (CDA) as, it “is a type of discourse analytical research that primarily studies the way social power abuse, dominance, and inequality are enacted, reproduced, and resisted by texts and talk in the social and political context” (Dijk, 1998, p. 352). It is concerned with the analysis of discursive practices such as hegemony, racial injustice and discrimination of basic rights within socio-political context.

Meyer (2001) defines the term ‘Critical’ in CDA as understanding of the language from its social, political and economic perspective. Through the analysis, it tries to trace the connections between the powers exerted through language within a wider socio-political context. According to him, the discourse reflects the ideology and identity which are constructed through the social relations. CDA not only traces the power and social relations but it also seeks to challenge and criticize it.

The techniques used by CDA are specified by Corson (1995) who says that critical discourse analysis applies certain techniques like textual and stylistic features in order to find out the social relations and identities from the linguistic perspective. Blommaert (2005), Jorgensen and Phillips (2002) further elaborate that linguistic and social variables in CDA research are closely connected to each other because “whenever we approach any text or discourse, we approach it from linguistic perspective and emphasize how language choice in a particular discourse pertaining the socio-political context tries to analyze the discursive practices” (Sipra, 2013, p. 29) .

Fowler (1996) says that the purpose of CDA is to facilitate the conscious rising of the people regarding any exploitative and manipulative discourse in its socio-cultural context. It aims at making transparent the connections between discourse practices, social practices, social structures, and language, connections that might be opaque to the layperson.

Along with CDA, another key methodological strand within this research is the use of corpus to locate the lexical items that indicate biased identity construction of non-white perpetrators. This construction of biased identity is because of the ‘us vs. them’ dichotomy as discussed by Said (1978). He emphasized that West uses its power to create European self-image. They have established a collective concept of ‘us’ against non-Europeans proving that their culture is rich and superior to rest of cultures of this world (Said, 1978). Westerns have created a hierarchy which is based on binary oppositions in which one is privileged and the other is underprivileged. These binary oppositions range from universal binaries such as white/black, literate/illiterate, male/female, light/dark to culturally weighed binaries like self/other (Childs & Fowler, 2006). The Self is always presented by positive terms whereas Other is always taken in a negative connotation (Childs & Fowler, 2006). The self is familiar and the other is strange (Said, 1978). Such dichotomies are also present in newspaper articles. They have been scrutinized by studying the linguistic items used for the representation of American and non-American shooters with the help of Corpus.

Corpus is defined as a collection of texts that are selected by following a defined and detailed linguistic criterion (Leech 1991, Sinclair 1996 & Williams 2003). It is “a

collection of naturally occurring examples of language, consisting of anything from a few sentences to a set of written texts or tape recordings, which have been collected for linguistic study” (Hudson, 2002, p. 2). Corpus Linguistics is a linguistic analysis and it is based on the idea that the meaning of “naturally occurring” discourse can be better understood if it is taken in its real world function and context (Vessey, 2013). Genre of discourse, variety of language and subject comprise the contents of a corpus, while its authenticity is based on corpus samples which are spoken or/and written texts (Nesselhauf, 2011, McEnery & Wilson, 1996).

CL can help in identifying the significance of including and excluding certain linguistic and textual features. It uses empirical evidence as a reliable source in order to classify and investigate the linguistic structures being used. These evidences from the data also help in interpreting the observed discourse to find out the “probabilities, trends, co-occurrences or groupings of features” (Cheng & Lam 2013, p.175). From these observations, the preferences and constraints that are associated with these features can be gauged “for what they are” (Hanks, 2012, p. 405). These conclusions are mostly based on the study of collocations, concordances, adjectives and keywords. These features represent real-world examples of language which can be analyzed qualitatively in order to understand their practical and attitudinal meaning (Sinclair 1996, McEnery & Wilson 1996). Therefore, Corpus Linguistics deals language as an inseparable part of ideology, such linguistic analysis provides a high level of objectivity which cannot be achieved through intuition and introspection (Baker, 2006, p. 3).

In studies related to media, the researchers have attested to the high degree of objectivity present in researches assisted by Corpus Linguistics. The cautious use of programs and statistical methods used in corpora can help in minimizing the influence of preconceived notions about semantic and linguistics content. It also provides a base for conducting a fixated linguistic study (Baker, 2006). Despite being two different methodological stances, CDA and corpus are used together quite often in researches. Hence a lot of scholarship is available on the combined use of these two approaches. A few of such references are mentioned below:

Corpus Linguistics (CL) and Critical Discourse Analysis (CDA) deal largely with quantitative and qualitative analyses respectively. The use of these approaches makes quantitative results more consistent and reliable, and qualitative analysis more specific and accurate (McEnery & Wilson, 2001). Corpus Linguistics (CL) uses a large collection of texts which are considered to be representative and this approach helps in keeping the researcher and data at comparatively more distance (Hoffmann et al., 2008). Yet this approach can also put the findings at risk since the statistical tools used in the research can give more generalized explanation of the data under study. On the other hand, CDA helps in analyzing the topic under study more profoundly. The data is analyzed by keeping in view the socio-political context. The findings of the study can be broadened by using these two approaches in a research. Using corpus linguistic tools in CDA can aid at three levels as explained by Partington (2003). On a simpler level, the techniques used in corpus can help in situating the ideas that were already being noted. The other level helps in uncovering the patterns which were not considered important previously. In

between these two levels, the intuition of researchers can be supported or revised by focusing on the grounds of their doubts and suspicions.

Corpus-assisted discourse study can contribute in close linguistic and textual study, analyzing social practices along with their relation to social structures, and analyzing the interpretations based on shared experiences of people (Fairclough, 1992; Cheng & Lam, 2012).

In the present research, the sample chosen to reveal the underhand propagation of racial stereotypes through mass media belongs to the coverage of mass shooting incidents in USA. This issue of mass shootings has a very contemporary relevance as they have increased in frequency in past few years. Hence a lot of research has been conducted to explore its causes and effects.

Gun violence has become a dilemma in the United States as shootings have been occurring at an alarming rate in form of individuals being targeted, gang violence or during robberies (Ross, 2009). Mass shooting has also become more frequent in past two decades in which victims are randomly targeted by the perpetrators. News organizations are creating a considerable amount of profit by presenting this news in such a way that it attracts a large number of audiences (Chermak, 1995). But the rate by which these perpetrators are represented by the media has also a great impact on establishing the stereotypes about the criminals. For instance, misrepresentation of a particular group as perpetrator may result in public anger towards that group by creating racial or ethnic stereotypes (Barlow, Barlow, & Chiricos, 1990). Studies also provide proof that racial minorities such as African Americans are mostly overrepresented in the news of any incident of violence or mass shooting (Frisby, 2015, Dixon & Linz, 2000). Television news stories have very rarely identified black suspects from their names; they are mostly represented as unemotional and self-possessed (Entman, 1990). If the victims are *blacks*, then such news are not given due coverage (Weiss & Chermak, 1998; Pritchard & Hughes, 1997). One study has also discussed that if the victims of the incident are *whites* than the articles are written on the sympathy of *whites* then *blacks* being victimized (Reiner, Livingstone & Allen, 2003). Chiricos and Eschholz (2002) found out that *blacks* and hispanics are most likely to be inappropriately portrayed as criminals as compared to the *Whites*.

As these studies show that media has not just given the coverage to the tragic incidents rather it has focused more on the identity of the criminals. Mainly it has manipulated the crime stories based on the racial conflicts. This study also focuses on investigating the type of coverage given to mass shooting incidents in American newspapers. Articles included in the study are based on American shooters and the others from different racial backgrounds. In order to identify the representation given to these shooters; words and adjectives related with the depiction of the shooters from the formulated corpora will be studied.

Language is considered as a mean of expressing something and writing is taken as a form of language (Smith, 1989). We need words to write down a text. Words in English

language are divided into word classes, for example verbs, nouns, adjectives, adverbs and others (Jonsson, 2007). Among these word classes some belong to lexical words as new words can be added with them, such as nouns, adjectives and words. While other classes belong to grammatical words as seldom or never words are added to them, for example pronouns (Barber, 2000). For this study adjectives are selected among other word classes as they are used for creating a clear image of the things being discussed in a written text. They are defined by Yule (1996) as “words used typically with nouns to provide more information about the ‘things’ referred to (happy people, large objects, cute creatures, stupid ideas)” (Yule, 1996, p. 88). Adjectives are used to explain the qualities, condition or a character (Sager & Svartvik, 2002). They are mostly found within written texts where they are commonly used to describe characters or settings (Jonsson, 2007). Moving from the general terminologies to the key components that are at play in the study, this section places the research in the existing body of knowledge in the relevant discipline along with the gap established for our research. This study examines following research questions

1. How does the use of lexical items in American news articles frame the ethnic background of the shooter?
2. How does the use of lexical items construct the identity of white American shooters?

3. Research Methodology

The present study falls under the realm of corpus based-mixed method research. An interpretivist approach and descriptive method was used for the analysis and interpretation of the data. It is a corpus-based study because it uses a deductive approach. The formulated corpora acted as a catalyst in order to support the pre-existing theoretical construct about media that it is not an innocent tool rather it is used as a main source of propagating and developing racial stereotypes (Weiss & Chermak, 1998).

This study is anchored in the concept of Critical Discourse Analysis put forth by (Fairclough, 1992). He defined CDA as a kind of analysis which brings to light various odd relationships between different social and ethnic groups. According to him, power is exercised through hegemonic control and dominant discourse of the elite (Fairclough, 1992). He focused on unfolding the role of language in the discriminating of social relations and the discursive practices which exploit the rights of masses. He is of the view that language serves as a tool for exercising power and hegemony. The present study focuses on the use of lexical items within American news articles in order to frame the ethnic background of the shooter. The description of White and Non-white perpetrators given in the American newspaper articles may reveal the biasness towards the people of different racial backgrounds. The analysis of certain lexical items can help in investigating the use of language in order to set stereotypes about the people who are considered as “Non-*whites*” because of their ethnicity. They are mostly considered as unemotional who are more inclined towards violence and terrorism. Such representation by the dominating group distorts the image of the Non-*whites* which results in their marginalization as “others”.

The study is also based on this very concept of “others” put forth by Edward Said. According to him East controls the West by misrepresenting it (Said, 1978). West is mostly represented by East according to Eurocentric perspective which gives it an inferior status. East is experiencing the dominancy over West and this concept of “otherness” is maintained by misrepresentation and marginalization. This study also focuses on scrutinizing the role of media aiding in constructing the image of Non-*whites* as criminals and terrorists. The descriptions are manipulated according to the racial background of the perpetrator, hence media is playing an important role in propagating the misconception about the people who are considered non-*whites* by the Americans.

The sample consisted of 32 articles related to sixteen major mass shooting incidents in America. 8 of these incidents were carried out by White American perpetrators and 8 by non-white naturalized American shooters. The names of the selected shooters are given in the table 1.

Table 1: List of mass shooters selected for the corpora

White American Shooters	Naturalized American Shooters
James Holmes	Sulejman Talovic
Nikolas Cruz	Omar Mateen
Adam Lanza	Rizwan Farook
Aaron Alexis	Seung Hai Cho
Devin Patrick Kelley	Jiverly Wong
Stephan Paddock	Mohammad Youssef Abdulazeez
Dimitrios Pagoutzis	Maj. Nidal Malik Hasan
Christopher Harper Mercer	One L. Goh

Two Corpora of 38,582 words in total were compiled based on these articles. One corpus is based on the white perpetrators and the other is based on Non-*Whites*. However, a norm count of 14,000 words per corpora were analyzed to avoid the discrepancies in the data. The articles were collected from the top American newspapers namely *The New York Times*, *The Guardian U.S Edition*, *USA today*, *The Washington Post* and *New York Daily News*. The list of these articles is provided in the appendix. These incidents were specifically chosen for several reasons. Firstly, the juxtaposition of these articles revealed the role of race in forming a biased attitude of media towards the non- white perpetrators. Secondly, all of these incidents occurred in the 2000s, giving it a more present-day relevance. Finally, these incidents were chosen because they captured a lot of media attention all over the world. They became a burning issue and were followed closely by the American public. The data was collected from newspaper articles in the form of words. Two corpora were formed from these articles from the scratch. These corpora were analyzed with the help of a concordance software program WordSmith tools 5.0 (WST). This software provides diverse corpus counts and numerical data; it also shows how words act in a text and shows this data in varied formats in order to state the features of linguistic study. There are three main applications of WordSmith tools, which are wordlist, concord and keywords. This study is carried out by using Wordlist and Concord. The Wordlist helped in obtaining the lists of the words which were displayed according to their frequency of occurrence within the corpus. Then with the help of

concord application, KWIC (Keyword in Context) or concordances of the adjectives were located which helped in providing the most recurrent collocational pattern of the adjectives. This helped in analyzing the use of adjectives in the newspaper articles. These adjectives from the corpora were extracted with the help of LancsBox32. This software provides the feature of “Whelk” where the data can be annotated for different parts of speech. It gave the list of adjectives along with their context. From this list, adjectives were selected manually which are used to under-represent or over-represent the shooter and the shooting incident in tables.

4. Data Analysis

Media is the most important and efficient tool to get the information about everything including whatever is happening around the world. The previous studies have already established the fact that the use of media is far from innocent and neutral (Baker, 2006). Rather it is an instrument to control the minds of the masses by propagating the dominant agenda. Carrying forward this assumption, the aim of this research was to analyse the newspaper coverage of American mass shootings and to investigate how the difference in the use of words and particularly adjectives constitute a racial project that results in reinforcing the disparity between white and non-white American community.

The corpora analysis substantiated the above assumption in several ways. Firstly, the use of adjectives with the non-white shooter/perpetrator reveals the implicit strategies used by the media to propagate and reinforce the racial or ethnic stereotypes. Most of the identified adjectives consisted of the ethnicity of the shooter. All of the top American newspaper identified the shooters as immigrants from various nations. Although most of them had been born and raised in America but their ethnic background was deemed an important factor in their inclination towards violent behaviour. Figure 1 indicates the frequencies with which the perpetrators had been identified with their ethnicities

Table 2: Frequencies of the identified ethnicities of the shooters

Perpetrator	Ethnicity	Frequency
Sulejman Talovic	Bosnian	18
Omar Mateen	Afghan	11
Rizwan Farook	Pakistani/Saudi	11
Seung Hai Cho	Korean	7
Jiverly Wong	Vietnamese	2
Mohammad Youssef Abdulazeez	Jordanian	4
Maj. Nidal Malik Hasan	Palestinian	2
One L. Goh	Korean	5

The frequent references to the ethnicities of the killers propagate the racial stereotypes which are then passively received by the audience. Race takes the centre stage in media and public discourse as the reaction of the media and eventually of the public to such incidents is determined by the race of the shooter.

However the exclusion of the race when the shooter is a white American man points towards white privilege and shows the tendency of media to take on a hegemonic white perspective. A KWIC search for the word white in the corpus of articles about white American shooters yields significant results that throw light on the dominant discourse of white supremacy as stated by Fairclough (1992). Figure 1 below shows the KWIC results for *white*.

In the figure 1, concordance lines 1, 6, 7, 8 and 11 refer to the white house where as lines 3 and 4 are references to some other white people (not the shooter). Concordance lines 2 and 9 describe the church and shorts. Line 10 refers to a white suspect. However the context of this line reveals a racial project in itself. Hence, in all the articles, only once a shooter has been described as a member of the white community. This lack of mention of the white race conveys the idea that the perpetrator acted alone and his actions are in no way representative of the whole white community. On the other hand, references to the ethnicity of non-white shooter generalizes the actions of one member as representative of the whole ethnic community. Thus, legitimizing the “Foreign” or “Other” status of such communities as proposed by Said (1978).

N	Concordance
1	ordered flags flown at half-staff at the White House and all federal buildings
2	First Baptist Church is predominantly white , and Mr. Kelley is white. The
3	predominantly white, and Mr. Kelley is white . The authorities said Mr. Kelley
4	attack was racial hatred — Mr. Roof, a white supremacist, plotted an assault
5	, he said, the suspect, described as a white male in his early 20s, was seen
6	he spoke about the shooting from the White House’s briefing room. “Our
7	Anacostia River a few miles from the White House and about a half-mile from
8	tossed firecrackers over the fence at the White House, causing loud bangs and
9	Service agents, who tackled a man in white shorts and a T-shirt on
10	two suspects, one said to have been a white man wearing a navy-style kahki
11	, a man was arrested outside the White House after throwing firecrackers

Figure 1: KWIC search results for white

Another way through which the racial project is quite evident is through the difference in use of particular terminologies that construct the identity of the shooter as well as guide the discourse around the shooting incidents. The most prominent expressions are the references to religion especially that of the perpetrators with Islamic background. If an incident whether a mass shooting or in any other form of violence is carried out by a Muslim, it is instantly labelled as an act of terrorism. It is not individualized rather generalized to the entire Muslim community or the religion Islam. However when similar kinds of incidents are carried out by white American men, they are hardly ever labelled as terrorism or Extremism. It may be noteworthy that although terrorism is defined by Oxford dictionary as an unlawful use of force or intimidation to gain political benefits. In no ways it specifies any religion or religious ideology. But the

dominant media discourse in America attaches words such as terrorism, extremism and radicalization with Islam and with perpetrators who are considered *foreign*.

Figure 2 throws light on the marked difference in use of words such as “terrorism/terrorist”, “extremism/extremist”, “radicalization/radical”, “religious” and “Islam/Muslim” in both the news articles about white American shooters and non-white American shooters.

Table 3: Frequencies of selected words in corpora

Rank	Term	Frequency in Corpus 1 (White American Shooters)	Frequency in Corpus 2 (Non-white American Shooters)
1.	Terrorist/Terrorism	2	18
2.	Extremism/ Extremist	0	6
3.	Radical/ Radicalized	0	12
4.	Religion (Christianity or Islam/Muslim)	0	16
5.	Religious	0	6

In the above mentioned figure1 the word terrorist or terrorism has been used many times while referring to the shooters or to the incidents carried out by *non-whites*. The statistics above entail that religion especially Islam is depicted as a major motive behind such mass shootings when the shooter is a Muslim. But the violence by white perpetrators is never labelled as terrorism nor as an act of extremism and neither is Christianity ever deemed a possible influencer in such scenarios. Figure 2 below sheds light on the ways in which non-white shooters are described and represented as terrorists in the selected articles through phrases which label them as radicals with quite certainty.



Figure 2: Description of non-white shooters in selected articles

The incidents and the shooters are often described through phrases such as ‘a terrorist or an extremist act’, ‘killings with terrorist roots’, ‘inspired or directed by a terrorist organization’, ‘with ties to an extremist ideology’ and ‘with latent extremist sympathies’. Such claims are presented as objective facts which are then passively taken up by the readers and contribute to fuelling the already existent hostility towards the foreigners and thus marginalizing them as “others”.

On the other hand, white perpetrators are almost always humanized by declaring mental health issues as the root of the problem. In all of the articles under study about white American shooters, each one of them refers to mental health as major driving force behind the incident. The frequent reference to mental issues generate sort of sympathy for the killers in the mind of the readers. Conversely, mental health issues are never an option in case of Non-white shooters. Figure 3 shows the occurrence of word “mental” in corpus of white American shooters’ articles.

unknown what contribution, if any, the shooter's	mental	health issues made to his attack... Those
health issues made to his attack... Those	mental	health professionals who saw him did not
mass shootings, the president focused on the	mental	health of the perpetrator and claimed that
a factor. "We have a lot of	mental	health problems in our country, as do
have been much worse." "This is a	mental	health problem at the highest level. It's
Asia, President Trump called the shooting a	mental	health problem at the highest level" and
campus. The suspect had reportedly been receiving treatment for	mental	health issues. Twelve of those killed have

Figure 3: KWIC concordances for word “mental” in Corpus 1

The manipulative nature of media is also evident through the difference in depiction of shooters and the weaving of facts around the shooting incidents. The analysis of the KWIC concordances has yielded that the articles about the white American perpetrators are mostly constituted of objective facts i.e. about the objective details of the incidents and the lives of the shooters. On the other hand, the articles about shooters with different ethnic background focus on the subjective details about the lives of the killers. These details include assumptions as well as opinions which are circulated as objective facts. Figure 4 and figure 5 below show the difference in the contexts around two mass murderers such as Omar Mateen and Adam Lanza

Lanza was obsessed with mass murder Adam	Lanza,	the 20-year-old who convulsed America when he shot
other students and a teacher in Colorado.	Lanza,	also kept a cutting from the New
report also illuminates the lengths to which	Lanza	went in planning the killings. GPS routes
it clear that there was no evidence	Lanza	had received formal medical treatment for any
which underlines the scale of violence that	Lanza	inflicted in a remarkably compressed time frame.
believed to be the shot with which	Lanza	killed himself, was heard at 9.40.03- just
bed, and others- were owned by Nancy	Lanza.	She took her son to shooting ranges,
happened to her. "One witness indicated that	Lanza,	did not have an emotional connection to
whom law enforcement sources identified as Adam	Lanza,	20. Police said that Lanza first apparently
as Adam Lanza, 20. Police said that	Lanza,	first apparently killed his mother, Nancy Lanza,
Lanza first apparently killed his mother, Nancy	Lanza,	at the home they shared in Newtown.
Hook Elementary, the school at which Nancy	Lanza,	taught kindergarten. In the car, authorities said,
Glock. At the school, authorities said, Adam	Lanza,	shot and killed six adults— including the
misidentified the shooter as Adam's brother Ryan	Lanza,	24. Based on reports from those sources,
many other news outlets also identified Ryan	Lanza,	as the suspected shooter. Ryan Lanza was
Ryan Lanza as the suspected shooter. Ryan	Lanza,	was taken into custody near his home
same street as the Lanzas, recalled Adam	Lanza,	as withdrawn, but not threatening in any
still something that troubled her about Adam	Lanza.	When Israel heard that Ryan Lanza— Adam's
Adam Lanza. When Israel heard that Ryan	Lanza—	Adam's brother— had been described by law
that he had a connection to the	Lanza	family: he had helped Nancy Lanza with
the Lanza family: he had helped Nancy	Lanza	with dead batteries over the years, and

Figure 4: KWIC concordance of *Lanza*

People who knew Orlando nightclub shooter Omar Mateen describe him as a man who had a lifetime of angst and embarrassment, Omar Mateen was on the verge of realizing a to fall apart. At a class barbecue, Mateen told a fellow cadet he was "allergic" pork, and he got teased about it. Mateen blew up, recalled several cadets who were Tech shooting that left 32 victims dead— Mateen asked a classmate whether he would report documents show. The next thing students knew, Mateen had been kicked out of the academy in documents as "at best extremely disturbing," Mateen was never charged, and so the incident in the worst shooting in U.S. history. Mateen appeared conflicted about his religion and his families would get together to share meals. Mateen and his three sisters stood out as do." But friends and neighbors recall that Mateen was a troubled child, often unpredictable, angry at Mariposa Elementary, where his mother taught Mateen in fourth and fifth grades. Winkler recalled taking other kids' toys and acting like a bully, especially toward girls. Winkler said that Mateen acted superior to others and that teachers an angry kid," said Winkler, who remembered Mateen having few friends. He was not sure Mateen was ever diagnosed with any learning difficulties about his son. On Sept. 11, 2001, Mateen was 14 and a sophomore at the school functions. Officials declined to prosecute, but Mateen later listed the incident on job applications students watching as the second plane hit. Mateen was smiling. It was almost like surreal tower get hit on a classroom TV, Mateen stood up and claimed that Osama bin right across the face." Other classmates described Mateen as disruptive, but some said he was clown than a troublemaker. Several said that Mateen, who was overweight, often got picked on. said he rode the school bus with Mateen, for several years. "He was a chubby But it didn't work." Those who knew Mateen then said they saw nothing to suggest Mateen's house and being surprised to see Mateen pull out a mat and pray. Addis a classmate at Martin County High, said Mateen played football for a short time and an extremist ideology and did not believe Mateen could have celebrated 9/11. "I know Omar

Figure 5: KWIC concordance of *Mateen*

The above two tables reflect the disparity in treatment of the two shooters. Most of the contexts surrounding Adam Lanza are somewhat objectively describing the shooter himself or the time or place of the incident. Conversely, the contexts surrounding Mateen mostly consist of opinions or assumptions about his childhood, his views or life in general. The manipulative role of media can also be substantiated by observing the adjectives which are being used in order to define the perpetrators or incident. Adjectives from the corpora were identified with the help of LancsBox. This software gave the list of adjectives along with their context so it helped in selecting the adjectives which were used to over-represent the Non-Whites as terrorist and to under-represent the Whites merely as criminals. The list of adjectives showed that the incident where Non-Whiter shooter is involved is described as more “terrifying”, “tragic”, “dark” and “explosive” and “worst” instance (Washington Post, June 17, 2016, New York Daily, April 3, 2012). It is evident that with the help of such explanation, media has tried to magnify the intensity of the incident. When these words were looked in the corpus based on the incidents related to White shooters, no such portrayal is given to the event. The following table shows the number of instances in which the mentioned adjectives are used in both corpora.

Table 4: differences in frequencies of selected adjectives in the two corpora

Adjective	Frequency in Corpus 1 (White American Shooters)	Frequency in Corpus 2 (Non- White American Shooters)
Dark	0	7
Terrifying	0	2
Tragic	1	3
Explosive	0	2
Worst	2	4

The above table also shows that the intensity of the events is masked where shooters are *whites*. This misrepresentation is not limited to the events only but it can also be observed that shooters are being judged on the basis of their ethnicities as discussed earlier. Through the study of adjectives it can also be noted that Non-White shooters are explained as more “abusive”, “rude”, “harsh”, “aggressive” and “crazy.” These attributes are given to them on the basis of the assumptions about their mentality and childhood.

They are shown to be more “religious”, “strict” and “conservative” beings showing “violent”, “irate” and “curt” behaviour (New York Times, April 04, 2009 & The Guardian, April 7, 2009, Washington Post, June 18, 2015). The table shows that these adjectives are used with *non-white* shooters and no such description is given of *white* shooters.

Table 5: Differences in frequencies of negative adjectives in the two corpora

Adjectives	Frequency in Corpus 1 (White American Shooters)	Frequency in Corpus 2 (Non-White American Shooters)
Angry	0	13
Mad	0	20
Disrespectful	0	2
Crazy	1	2
Obsessed	0	1
Abusive	0	2
Rude	0	2
Violent	3	4
Disturbed	0	2
Awful	0	1
Murderous	0	1

On the other hand, when the perpetrator is White, his act is justified by considering it as a “psychotic episode”. If his childhood is discussed, he happened to be “shy” and “quite” (The Washington Post, September 28, 2017). The act of shooting of *Whites* is portrayed as a result of their “mental-health” which caused them to show “anger-fuelled” behaviour (The Guardian, November 25, 2013). With the help of these adjectives, shooters are under-represented as criminals but rather as mentally “troubled” patient who is in need of “medical” help (The Guardian, October 2, 2015). Following table helps to compare the misrepresentation of both shooters by under-representing the white shooters.

Table 6: differences in frequencies of health-related adjectives in the two corpora

Adjectives	Frequency in Corpus 1 (White American Shooters)	Frequency in Corpus 2 (Non-White American Shooters)
Anger-fuelled	1	0
Mental-health	5	0
Mental	20	3
Medical	7	1
Psychotic-episode	1	0
Troubled	3	0

5. Conclusion

This study investigated the manipulative role of media in developing and circulating racial and ethnic stereotypes. A corpus based approach was taken to analyze the newspaper coverage of American mass shootings. A comparison was done of the words and adjectives used in news articles when the perpetrator is of white descent and when he is of any other ethnic background. For this two corpora of 38,582 words in total were formulated based on sixteen mass shooting incidents. Among them eight incidents were carried out by white shooters and eight by non-white shooters. The corpora were formulated from news articles taken from top American newspaper agencies such as The Guardian U.S Edition, The New York Times, New York Daily and The Washington Post. The data was analyzed in the light of CDA approach presented by Fairclough and the concept of *other* put forth by Edward Said.

The data analysis generated many ways in which newspaper industry constitutes a racial project when it comes to non-white people. The frequency with which the media represents the alleged perpetrator has a significant impact on the development or circulation of stereotypes. These labels serve as a major driving force behind formation of certain attitudes towards people of various ethnic backgrounds. The findings revealed that a non-white shooter is almost always described through his ethnicity by mention of adjectives such as Korean, Afghan, Bosnian, Saudi and Vietnamese. Whereas, a deliberate omission of the race is found when the perpetrator is a white man. This entails that a crime committed by a non-white is generalized to his entire ethnic community but one committed by a white person is considered as a lone attack, one not representative of the white race. The findings further revealed the disparity and double standard in use of terms such as terrorism, extremism and radicalization in description of mass shooting incidents. “Terrorist”, “extremist” and “radical” is mostly used with Muslim perpetrators and with perpetrators of immigrant status to some extent but never with white mass murderers. These white killers are justified or humanized by declaring mental health issues as the root of the problem. The racial or ethnic stereotypes are also propagated by presenting subjective opinions or assumptions as objective facts in case of *non-white* shooters. Their whole life particularly details surrounding their ethnic backgrounds are explored to determine the aggravating circumstances that might have caused the shooting. Conversely, stories about white shooters are much more likely to only include objective facts, such as the time, date and place of the shooting. The shooters and incidents are under-represented in the article through limited use of adjectives with negative connotations. But non-white shooters are over-represented and the severity of the incidents are magnified through extensive use of negative adjectives. This research aimed at familiarizing the readers with a possible role of American newspaper media in establishing, reiterating and propagating the stereotypes against naturalized American with different ethnic backgrounds as it is clear from the above study of the recent mass shooting incidents’ coverage by such top American newspapers. However, the future researchers may also investigate the manifestations of this disparate treatment of non-white Americans in different walks of life as well as the rise of *othering* after such mass shooting incidents. Researchers could also investigate the construction and propagation of “Multi-culturism’ in newspaper media as a possible root for homegrown terrorism.

References

- Barber, C. (2000). *The English Language, A Historical Introduction*. Cambridge: University Press.
- Baker, P. (2006). *Using Corpora In Disclosure Analysis*. London: Continuum.
- Barlow, M. H., Barlow, D. E., & Chiricos, T. G. (1990). Economic conditions and Ideologies of crime in the Media: A Content Analysis of Crime News. *Crime and Delinquency*, 41(1), 3–19.
- Blommaert, J. (2005). *Discourse: A Critical Introduction*. UK: Cambridge University Press.
- Corson, D. (1995). *Discourse and Power in Educational Organizations*. Toronto: OISE Press.
- Cheng, W., & Lam, P. W. Y. (2012). Western Perceptions of Hong Kong Ten Years On: A Corpus-driven Critical Discourse Study. *Applied Linguistics*, 34(2), 173-190.
- Chermak, S. M. (1995). *Victims in the news: Crime and the American news media*. Boulder, CO: Westview.
- Childs, P. & Fowler, R. (2006). *The Routledge Dictionary of Literary Terms*. New York: Routledge.
- Chiricos, T., Eschholz, S., & Gertz, M. (1997). Crime news and fear of crime: Toward an identification of audience effects. *Social Problems*, 44, 324-357.
- Chuang, A. (2012). Representations of foreign versus (Asian) American identity in a mass-shooting case: Newspaper coverage of the 2009 Binghamton massacre. *Journalism & Mass Communication Quarterly*, 89(2), 244-260.
- DeFoster, R. (2015). *Terrorizing the Masses: Identity, Mass Shootings, and the Media Construction of "Terror"* (Doctoral dissertation, University of Minnesota).
- Dixon, T. L. & Linz, D. (2000a). Race and the Misrepresentation of Victimization on Local Television News. *Communication Research*, 27, 547–573.
- Entman, R. M. (1990) Modern racism and the images of *Blacks* in local television news. *Critical Studies in Mass Communication* 7, 332-345.
- Fairclough, N. (1992). *Discourse and Social Change*. Cambridge: Polity Press.
- Fowler, R. (1996). 'On Critical Linguistics ': in Caldas-Coulthard, C. R. and Coulthard, M.
- Frisby, C. M. (2017). Misrepresentations of Lone Shooters: The disparate treatment of Muslim, African American, Hispanic, Asian, and White perpetrators in the US news media. *Advances in Journalism and Communication*, 5(02), 162.
- Hanks, P. (2012). The corpus revolution in lexicography. *International Journal of Lexicography*, 25(4), 398-436.
- Hoffmann, S., Evert, S., Smith, N., Lee, D., & Prytz, Y. B. (2008). *Corpus Linguistics with BNCWeb – a Practical Guide*. Frankfurt: Peter Lang.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Jonsson, C. (2007). *Adjectives make a better text, don't they?* (Unpublished doctoral dissertation). Högskolan Kristianstad, Sweden.
- Jorgensen, M. W. & Phillips, L. J. (2002). *Discourse Analysis as Theory and Method*. London: SAGE Publications Ltd.

- Leech, G. (1991). The state of the art in corpus linguistics. In K. Aijmer & B. Altenberg (Eds.), *English Corpus Linguistics* (pp. 8-29). London: Longman.
- Meyer, M. (2001). *Between theory, method, and politics: Positioning of the approaches to CDA*. In R. Wodak & M. Meyer (Ed.), *Methods of Critical Discourse Analysis*.
- McEnery, T., & Wilson, A. (1996). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- McEnery, T., & Wilson, A. (2001). *Corpus linguistics*. Edinburgh: Edinburgh University Press.
- Mingus, W., & Zopf, B. (2010). White means never having to say you're sorry the racial project in explaining mass shootings. *Social Thought & Research*, 57-77.
- Nawaz, S., Bilal, A. H., Kalsoom, M., Fayyaz, Z., & Nayyar, H. (2013). Media discourse and their implicit ideologies. *Asian Journal of Social Science and Humanities*, 2(2), 328-337.
- Nesselhauf, N. (2011). *Corpus Linguistics: A Practical Introduction* [PDF]. Retrieved 29 April 2015, from <http://www.as.uni-heidelberg.de/personen/Nesselhauf/files/CorpusLinguisticsPracticalIntroduction.pdf>
- Partington, A. (2003). *The linguistics of political argument: The spin-doctor and the wolf-pack at the White House*. London and New York: Routledge.
- Pritchard, D., & Hughes, K. D. (1997). Patterns of deviance in crime news. *Journal of Communication*, 47, 49-67.
- Reiner, R., Livingstone, S., & Allen, J. (2003). From Law and Order to Lynch Mobs: Crime News since the Second World War” In P. Mason, (ed) *Criminal Visions: Media Representations of Crime and Justice*, 13-32. Portland: Willan Publishing.
- Ross, S. (2009). Gun violence shows now signs of abating, Retrieved November 25, 2009 (<http://pubrecord.org/nation/6072/violence-shows-signs-abating/>).
- Sager, O. & Svartvik, J. (2002). *Engelsk Universitetsgrammatik*. Uppsala: Almqvist och Wiksell Tryckeri.
- Said, E. W. (1978). *Orientalism*. New York: Penguin.
- Sinclair, J. (1996). The search for units of meaning. *TEXTUS*, 9(1), 75-106.
- Sipra, M. A. & Rashid, A. (2013). Critical discourse analysis of Martin Luther King’s speech in socio-political perspective. *Advances in Language and Literary Studies*, 4, 27-33.
- Smith, F. (1989). *Writing and the writer*. UK: Biddles Ltd.
- Van Dijk, T. A. (1995). Power and the news media. *Political Communication and Action*, 6(1), 9-36.
- Van Dijk, T. A. (1998). *Ideology: A Multidisciplinary Approach*. London: SAGE publications Ltd.
- Vessey, R. (2013). Challenges in cross-linguistic and corpus-assisted discourse studies. *Corpora*, 8(1), 1-26.
- Weiss, A., & Chermak, S. M. (1998) The News Value of African American Victims: An Examination of the media’s presentation of Homicide. *Journal of Crime and Justice*, 21, 71-88.
- Williams, G. (2003). From meaning to words and back: Corpus linguistics and specialized lexicography. *ASp*, 91-106.
- Yule, G. (1996). *The Study of Language*. Cambridge: University Press.

Gender-based Study of Stance Features in Blogs through Corpus Techniques

Muhammad Yousaf¹
National University of Modern Languages, Islamabad

Ijaz Ali Khan²
Govt. College Lahore, Sawabi

Abstract

The current study explores the uses of stance features in blogs: Self mention, hedges, boosters and attitude markers as used by Pakistani male and female bloggers. A quantitative approach is used because of the descriptive nature of this study. The paper aims at capturing the differences or similarities in using stance features by male and female Pakistani bloggers. Hyland's (2005) interaction model is used to analyze the corpus which was compiled specifically for this study. The corpus consists of 20 blogs, 10 each by male and female bloggers. The findings of the study indicate that the female bloggers use a large number of first person pronouns than their male counterpart in their blogs. The findings further reveal that the female bloggers use hedges more frequently as compared to male bloggers. However, the boosters are more frequently used by male bloggers. Finally, it is concluded that the female bloggers prefer to use stance features which are related to uncertainty, whereas male bloggers use the features which are neutral in their tone but assert their points of view.

Keywords: attitude markers, boosters, hedges, self-mention, stance

1. Introduction

The study of gender and language not only provides a descriptive account of male and female discourse but also reveals how language functions as a symbolic recourse to create personal, social and cultural meanings and identities (Tannen, 2001). The issue of gender and discourse is deemed as an interdisciplinary endeavor shared by scholars in linguistics, anthropology, speech communication, education and literature. By and large, researchers tend to identify the distinctive social and cultural meanings as well as the constructed identities of the individuals in their use of a language (Kaplan, 1997).

Gender and discourse as a sub-field emerged in mid-seventies, which was in reality, a continuation of the feminist movement of 1970. The publication by Lakoff, Keythone and Henley (1975) is considered a ground-breaking event in the field of gender and discourse. As mentioned earlier, the pioneers took the issue from the aspect of feminism (Lakoff, 1975; O'Barr & Atkin, 1980; West & Zemmerman, 1975; Fisherman, 1983; Brown, 1980). These studies were based on inequality between two genders. These

studies also focused on role of women that is usually constructed through society.

Newman et.al (2008) conducted an empirical investigation where he cited a few researchers who worked on language and gender and found clear differences in the language used by men and women. The study by Newman et.al (2008) suggests that women tend to allocate comparatively a larger number of words to describe the psychological and social processes, whereas, men in greater detail describe object properties and impersonal topics.

To find a space for the present study, it would be more appropriate to take an aspect of the usual male and female discourse and analyze that in accordance to the context of the sampled occurrences. Generally, both male and female change their constructed role and identity (McElhinny, 1992) in institutional settings; they do not want to be seen unprofessional. The corpus consisting of spontaneous data is more authentic and reliable than a corpus that consists of formal speeches and well composed essays that had undergone repeated revisions. Hyland (2005) model of Stance and Engagement with its three elements of stance (i.e. self-mention, hedges and boosters and attitude markers) was used to study the data. The data was analyzed by using AntConc, a tool to study frequency count, collocation and concordance of the selected items.

2. Literature Review

Scholars and social scientists agree that the meaning of gender is culturally mediated and interactionally established and reinstated (Goffman, 1976). Tannen (1994) claims that gender and discourse have an indirect relationship. He further elaborates that discourse and gender are sex- cum-class linked rather than sex linked. The ways of speaking are not identified with any individual man or woman but are more associated with the class of women or the class of men in a given society (Schiffrin, 2001). Individuals, traditionally, align themselves with one or the other sex-class by their discursive practices, being part of their socially constructed identities. Ochs (1992) also endorses this view by saying that ways of speaking are associated with stances of women or men in a particular culture.

Feminism as a movement gave rise to the issue of gender related discourse in mid-seventies (see Schiffrin, 2001). Language and Woman's Place by Robin Lakoff (1975), Language and Sex: Difference and Dominance, by Thorne and Henley (1975) and Male and Female Language by Key (1975) were some of the works done in this regard. The writers highlighted the differences between men and women's speech, by analyzing women's speech as well as identifying the role of language in creating social inequality. Unequal gender relation has been a point of interest for the researchers (see West& Zimmerman, 1983). Brown (1980) found that Tenjapan women were qualitatively politer than men and emphasized in-group solidarity in their language use. Tannen (1990) in her "You Just Don't Understand", claimed that conversation between women and men could be taken as inter-cultural communication. Researchers (see Bucholtz, 1999; Sciffrin et. al, 2001; Boxer & Cortes, 1997) in general, focus on studying dominance and masculinity in language use in gender and language research.

The researchers working on feministic issues expanded their studies to multiple aspects of gender and discourse. Other social identities like ethnicity and social class were focused more in terms of social interaction (Bucholtz, 1999). This trend led to another major development, popularly known as dominance and masculinity in language (see Schiffrin, 2001). Tannen (1994) identified a specific pattern in a wide range of context, that men tend to discursively take up the role of expertise or authority. Coates (1997) proved the same through their individual studies (Schiffrin, 2001).

Newman et.al (2008) suggested several things to be considered before conducting a study. Firstly, the size of the data should be reasonably large. Secondly, there should be a wide range of contexts and settings to make the results more reliable and thirdly non-traditional methods should be utilized. The corpus of the current study consists of 20 blogs by Pakistani blog writer, 10 by each male and female blogger. To make corpus more reliable, it was ensured that no two blogs were written by one writer, hence, only one blog of a writer was included for the study so that individual differences could be minimized, and the general features of male and female bloggers could be studied. The studies conducted by Newman et.al (2008), Kaplan (1998), Martin (2001), Haufiku & Kangira, (2017), Hasani and Farahani (2014) adopted the same course which they claim proved fruitful. The method they used consisted of both quantitative and qualitative tools of analysis. The previous researches relied heavily on speech or written discourse but none of them has gone for the writing such as blogs. Hence, the current study has selected blogs as site for exploring gender differences in language. The research question of the study is:

1. What are the differences and similarities between male and female blog writers in terms of their use of self-mention, hedges and boosters, and attitude markers?

3. Research Methodology

For this study, a small size corpus, of 7968 words, was compiled. Keeping in view the scope of the study, only twenty blogs, ten each by male and female bloggers, were selected. Although the sampling was random, but the researchers strictly followed the following criteria for the selection of a blog:

- The blog must be written by Pakistani bloggers so that there is no variation in language because of cultural difference.
- The blog must be dealing with the social and political issues of Pakistan so that domain of language use remains the same across the blogs.

Mixed method approach was adopted which consisted of both quantitative and qualitative approach of data analysis. Quantitatively the data was analyzed through AntConc which an automated process was completely. Secondly, the obtained results were verified through concordance where qualitative human judgment was incorporated to judge each individual linguistic unit of self-mention, hedges, boosters and attitude markers. The result obtained was then put into tables specifically designed for this purpose and later on analyzed and discussed accordingly.

3.1 Theoretical Model

Hyland’s (2005) model of interaction was used as the theoretical framework for the analysis of the data. The model consists of two parts, which are stance and engagement. In the current study only the features of stance, which are self-mention, hedges, boosters and attitude markers, have been explored. Moreover, it is pertinent to mention here that Hyland’s (2005) model of interaction was based on interaction in academic discourse but the current study has extended its use to study blogs as well. In addition, the researchers of this study adapted Haufiku1 and Jkangira’s (2017) elements of hedges and boosters. The contours of the model of interaction are as follows:

Table 1: Hyland’s (2005) Model of Interaction

Interaction								
Stance				Engagement				
Hedges	Boosters	Attitude Markers	Self-mention	Reader Pronouns	Directives	Questions	Shared knowledge	Personal asides

4. Results and Discussion

This part of the paper gives a complete illustration of the data collected ranging from word token and types to analyses of the self-mention with its various linguistic manifestations such as I, me, my, mine, we, our, and us. The bloggers selected are all Pakistanis, freelance writers, commenting upon contemporary issues. The data has been analyzed quantitatively finding the total number of tokens, word types, and frequency count of each item belonging to self-mention.

4.1 Word Tokens and Word Types

The following table presents the results of the word tokens and word types present in the corpus of blogs which consists of equal number of blogs by male and female bloggers.

Table 2: Word Tokens and Types

	Male	Female
Word Tokens	9277	7968
Word Types	2599	1975

Table 2 shows total number of word tokens along the number of word types. It is obvious that almost equal size of texts has been selected for the study corpus to provide satisfactory answers to the questions put up above. The number of word token is 9277 for male bloggers and 7968 for female bloggers. This shows that the male bloggers write longer blogs as compared to female bloggers. Similarly, their variety of vocabulary use is also higher as compared to the female bloggers which is apparent from the word types.

4.2 Use of Self-Mention

The use of first person pronouns, both singular and plurals, are the best indicators to find out self-mention by the writer in his/her writing. Table 3 shows the self-mention by the male and female bloggers.

Table 3: Self-Mention in the Blogs

		Categories of Male Bloggers			Female Bloggers		
		Total	Used	Not used	Total	Used	Not used
Singular	I	84	09	01	103	08	02
	Me	14	03	07	24	04	06
	My	40	07	03	73	08	02
	Mine	01	01	09	01	01	09
Total		139	-----		201	-----	
Plural	We	42	07	03	83	08	02
	Us	11	04	06	20	06	04
	Our/ours	39	07	03	54	08	02
Total		92	-----		157	-----	
Grand Total		231	-----		358	-----	

The findings shown in the table 3 are indicative of the linguistic behavior of the people concerned. The usual relationship between language and gender has been entailed by giving a comparison between male and female blog writers in terms of the relative use of self-mentioning linguistic items. Each item was analyzed individually for all the writers and their values were put in the given table. Similarly, the sub-categories of self-mention were divided first into singular, having four items- *I*, *me*, *my*, *mine*, and plural having three items, *we*, *us*, *our/ours*.

Pronoun *I* has been used most frequently by the female bloggers. The relative use of this item in terms of the frequency count is highest, i.e. 103 against 84, on the part of female writers than male writers. So female writers have made enough use of the first-person pronoun which according to Schiffrin (2001) is indicative of a number of interpretations. Firstly, it shows assertive attitude on the part of the female writers and this is against the hypothesized result but at the same time it also indicates independent nature as against in-group (Schiffrin, 2001) feeling and a changed outlook of the female writers. Some of male writers have also used this first-person pronoun more frequently but that according to Newman et.al (2007) is not an unexpected behavior. As far as female blog writers are concerned, they have constructed another identity (see examples below) which is more or less context oriented rather than gender oriented. So, the evidence in the table entails that language variation is context oriented rather than gender oriented.

The use of *me* and *my* as instances of self-mention were also more in number by female blog writers (*me*=24 vs 14 and *my*=73 vs 43). It again shows the same stance-

oriented mode of language use by females. On the contrary, male writers' style is more engagement oriented (Hyland, 1997) as against female writers. This proves an answer to the question raised that whether language variation is gender oriented or context oriented.

The use of *mine* is less used (01 vs 01) in both male and female writings which may be due to its informal nature. The sampled writers from both the genders have tried to avoid slang or informal kind of expressions in their blogs, which was not expected in the beginning of this study. If we look at the use of first person singular pronouns (*I, my, mine*), the female writers have used *I* and *my* more frequently and higher in number than male writers. The overall frequency of 1st person pronouns is 201 times in the corpus by female writers against 139 times used in the corpus by male writers. In continuation to the findings mentioned above, female writers have used more expressions like *we* than male writers. The use of *we* in blogs of female writers is 83 times against 42 time used by male bloggers.

Similarly, the use of expression *us* is also found more in female writers' blogs than male bloggers. Female writers have used *us* 20 times and the male writers have used 11 times. This difference is not very large which is unlike the use of first person singular pronoun. Moreover, the use of last item in the self-mention, *our/ours*, is also higher in the blogs of female bloggers as they have used 54 times against the 39 times used by the male bloggers. Similarly, if we look at the first person plural pronouns (*we, us, our/ours*) that is also used in higher frequency by female writers than the male writers. The overall number of first person plural pronoun is 157 times used by the female writers and 92 times used by the male writers in the corpus of the Pakistani blogs. This phenomenon is indication of the same in-group feeling or engagement-oriented style of writing (Hyland, 1997), but here in Pakistani blogs it is used by females rather than male writers (see examples below). This kind of feeling was expected from the female writers as far as the gender is concerned. So, this is a repetition of the statement that language variation is context oriented.

4.2.2 Examples

The following are the examples by the blogs by female bloggers using higher self-mention than male bloggers.

"I am not complaining, *I* am showing the people who say what was the use of making Pakistan?"

"...to *me* it was intolerable.

"*I* turned home, called *my* people and told them exactly what *I* am about to say;....."

"If *we* live- *we* live, if *we* die-*we* die. *We* love Quid e Azam. *We* love Allama Iqbal....."

"..... adds up over time and *our* mood, *our* self-esteem and even on....."

4.3 Use of Hedges

In pragmatics hedging is a mitigating word or construction to lessen the impact of an utterance on the interaction between the speaker and the listener because of certain constraints, such as cultural, social, situational or relational. These are also used as

politeness strategies. Contrary to this view Hyland (2005) states that hedges “imply that a statement is based on plausible reasoning rather than certain knowledge”. Table 4 below shows the use of hedges by male and female Pakistani bloggers.

Table 4: Uses of Hedges by Bloggers

Hedges	Male Bloggers	Female Bloggers
Probable/probably	-	04
Indicate(s)	-	-
Believe	-	02
Speculate	-	-
Possible/possibly	-	02
Likely	01	01
Assume	-	01
Might	02	02
Could	10	11
Appear(s)	01	-
Seem(s)	03	03
May	07	05
Suggest(s)	02	01
Total	26	32

Table 4 indicates the tentativeness and the level certainty regarding the claims made by the bloggers. The number of hedges used by the female bloggers is 32 against those of male bloggers who used it 26 times. The frequency of hedges like *probably*, *could*, and *may* are the most attended instance used by the writers. The difference in the use of these terms of boosters is not that statistically significant but still shows a tilt towards the female bloggers and enough to highlight their assumed down toning behavior as compared to their counterpart bloggers. Here again the female bloggers made relatively more use of these hedges except for *may*, which were also referred to as ‘epistemic modalities’ by Martin (2001).

Through using these kinds of hedges, the writers tend to display their knowledge or information, avoid the face threatening acts on the part of the readers. Hence, the frequent use of the hedges by the female bloggers confirms the gender-based assumption of language appropriation (Hyland, 1998).

4.4 Use of Boosters

Boosters are the verbal constructions which are used to assert the point of view and add certainty in the statement given by the speaker. According to Hyland (2005), boosters are words which allow writers to express their certainty and help them mark their involvement with the topic and solidarity with the audience. The table 5 below gives the detailed view of the boosters used by the male and female bloggers.

Table 5: Uses of Boosters by Bloggers

Boosters	Male Bloggers	Female Bloggers
Certain/certainly	01	02
Definite/definitely	01	-
Clear/clearly	02	03
Obvious/obviously	-	01
Substantially	-	-
Show(s)	01	03
demonstrate	-	01
Highly	-	-
Should	19	04
Must	17	01
Total	41	15

The number of boosters used by the bloggers, as mentioned in table 5 indicate the rhetorical strategies employed by the writers in order to get their claims accepted. It is obvious from the data given above that the male bloggers have enough use of boosters as up to 41 times against 15 times by the female writers. The use of *must* and *should* is more abundant than any of the terms used or those which could be used to boost their claims. The terms *should* and *must* were used 19 and 17 times respectively by the male bloggers. Contrary to these, the female bloggers made the minimum possible use of *should* and *must* as up to 4 and 01 time respectively. The difference in the employment of the other instances of boosters is not very significant. The results shown clearly indicate the male dominant behavior in forwarding their claims which tends to be in conformity with the assumption regarding gender and linguistic attitude as argued by Hyland (1998). This is usually done when the writers tend to force their claim upon the expected audience without having substantial evidential support for their arguments. Moreover, these stress the shared information, group membership, and engagement with the audience (Hyland, 1999).

4.5 Use of Attitude Markers

“Attitude markers indicate the writer’s affective, rather than epistemic, attitude to Propositions” Hyland (2005). The table below shows the use of attitude markers by male and female bloggers.

Table 6: Attitude markers used by male and female bloggers

Attitude Markers	Male Bloggers	Female Bloggers
Agree/agreeable	-	-
Prefer(able/ably)	-	-
Fortunately/unfortunate	02	02
Hopefully	-	-
Appropriate/appropriately	-	01
Logical/logically	01	-
Remarkable/remarkably	-	-
Total	03	03

The given table 6 indicates the respective use of attitude markers by both male and female bloggers. The difference in even the use of such phrases is not that significant. Both the genders are rather equal in their preference for the attitude markers which tends to be affective rather than epistemic. As the table 6 shows that the whole corpus has only 3 instances of attitude markers each for male and female bloggers, so this shows that both genders remain neutral on the topic and theme of their blog which might help them increase their readership.

5. Conclusion

The present study leads us to answer the questions about the differences and similarities in terms of language use between male and female blog writers. These questions have been answered through taking the aspect of self-mentioning, hedges, boosters and attitude markers for the proposed comparison. A clear-cut difference has been found in terms of the language use of both the genders. The instances of self-mentioning pronouns were found with well-defined pattern by the female bloggers, and the male bloggers used lesser self-mention. Pronoun related to self-mention were used more in number by female blog writers which denote a specific kind of attitude on the part of female writers that is against the hypothesis in our culture. Similarly, hedges are also used more by female writers rather than male writers which show that they write in toned down, and euphemized language and appear to be more tentative and hesitant in making their point, whereas male writers are more direct and authoritative. This can also be seen using boosters which have been used almost three times higher by male bloggers than the female bloggers. It shows self-dependence and assertive stance by male bloggers. On the other hand, the male bloggers are highly influenced by the context they are work in, while using self-mention, as they try to be neutral by showing their lesser presence in their blogs. The use of attitude markers is very low by both male and female writers which highlights that the bloggers are impartial in their stance. There is a conflict between the patterns noticed by Schiffrin (2001) and the current study. Therefore, it is stated that bloggers from Pakistan use attitude markers in their own style.

References

- Aries, E. J., & Johnson, F. L. (1983). Close friendship in adulthood: Conversational content between same-sex friends. *Sex roles*, 9(12), 1183-1196.
- Bucholtz, M. (1999). "Why be normal?": Language and identity practices in a community of nerd girls. *Language in society*, 28(2), 203-223.
- Bucholtz, M. (1999). You da man: Narrating the racial other in the production of white masculinity. *Journal of sociolinguistics*, 3(4), 443-460.
- Boxer, D., & Cortés-Conde, F. (1997). From bonding to biting: Conversational joking and identity display. *Journal of Pragmatics*, 27(3), 275-294.
- Biber, D., Douglas, B., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.
- Brown, P. (1980). How and why are women more polite: Some evidence from a Mayan community. In *Women and language in literature and society* (111-136). Praeger.
- Brown, M. R., Allison, D. G., & Gilbert, P. (1988). Resistance of bacterial biofilms to antibiotics a growth-rate related effect?. *Journal of Antimicrobial Chemotherapy*, 22(6), 777-780.
- Coates, J. (1989). Gossip revisited: Language in all-female groups. *Women in their speech communities*, 94-122.
- Coates, J. (1997). One-at-a-time: The organization of men's talk. *Language and masculinity*, 107-129.
- Dovidio, J. F., Brown, C. E., Heltman, K., Ellyson, S. L., & Keating, C. F. (1988). Power displays between women and men in discussions of gender-linked tasks: A multichannel study. *Journal of personality and Social Psychology*, 55(4), 580.
- Goffman, E. (1976). Gender display. In *Gender advertisements* (pp. 1-9). Palgrave, London.
- Hassani, M. T., & Dastjani Farahani, M. (2014). A discourse analysis of gender differences in the use of hedging devices in applied linguistics research articles. *English Language Teaching*, 1(1), 59-73.
- Haufiku, N. K., & Kangira, J. (2017). An Exploration of Hedging and Boosting Devices Used in Academic Discourse Focusing on English Theses at the University of Namibia. *Studies in English Language Teaching*, 6(1), 1.
- Hawton, K., Rodham, K., Evans, E., & Weatherall, R. (2002). Deliberate self harm in adolescents: self report survey in schools in England. *Bmj*, 325(7374), 1207-1211.
- Hyland, K. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of second language writing*, 6(2), 183-205.
- Hyland, K. (1998). Hedging in scientific research articles (Vol. 54). John Benjamins Publishing.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse studies*, 7(2), 173-192.
- Kaplan, A. (1998). Manifest domesticity. *American literature*, 70(3), 581-606.
- Kaplan, R. B., & Baldauf, R. B. (1997). Language planning from practice to theory (Vol. 108). *Multilingual Matters*.
- Key, M. (1975). Male/female language. *New York*.
- Kendall, S., & Tannen, D. (2001). Discourse and gender Text. *The Handbook of Discourse Analysis*. Oxford: Blackwell, 46-63.

- Lakoff, R. (1973). Language and woman's place. *Language in society*, 2(1), 45-79.
- Lakoff, G. (1975). Hedges: A study in meaning criteria and the logic of fuzzy concepts. In contemporary Research in Philosophical Logic and Linguistic semantics (pp. 221-271). Springer, Dordrecht.
- Martin, E. (2001). *The woman in the body: A cultural analysis of reproduction*. Beacon Press.
- McElhinny, B. (1994). An economy of affect. Dislocating masculinity: *Comparative ethnographies*, 158-170.
- Miller, R. J. (1983). How many traps should a crab fisherman fish?. *North American Journal of Fisheries Management*, 3(1), 1-8.
- Mulac, A., Studley, L. B., & Blau, S. (1990). The gender-linked language effect in primary and secondary students' impromptu essays. *Sex Roles*, 23(9-10), 439-470.
- Mulac, A., Jansma, L., & Linz, D. (2002). Men's behavior toward women after viewing sexually-explicit films: Degradation makes a difference. *Communication Monographs*, 69(4), 311-328
- Newman, L. (2007). *The Cambridge companion to Locke's "Essay concerning human understanding"*.
- Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*, 45(3), 211-236..
- O'Barr, W. M., Atkins, B. K., McConnell-Ginet, S., Borker, R., & Furman, N. (1980). *Women and language in literature and society*.
- Ochs, E. (1992). 14 Indexing gender. Rethinking context: Language as an interactive phenomenon, 11, 335.
- Schieffelin, B. B., & Ochs, E. (Eds.). (1986). *Language socialization across cultures* (No. 3). Cambridge University Press.
- Schiffrin, D. (2001). Language and public memorial: America's concentration camps'. *Discourse & Society*, 12(4), 505-534.
- Thompson, S. A., & Mulac, A. (1991). A quantitative perspective on the grammaticization of epistemic parentheticals in English. *Approaches to grammaticalization*, 2, 313-329.
- Thomson, R., & Murachver, T. (2001). Predicting gender from electronic discourse. *British Journal of Social Psychology*, 40(2), 193-208.
- Tannen, D. (1990). Rethinking power and solidarity in gender and dominance. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 16, No. 1, pp. 519-529).
- Tannen, D. (1994). *Gender and discourse*. Oxford University Press.
- Tannen, D. (2001). Discourse analysis. *Linguistic Society of America*. Retrieved from http://www.philsci.univ.kiev.ua/UKR/courses/asp/asp-lit/tannen_d_hamilton_h_e_schiffrin_d_eds_the_handbook_of_discou.pdf
- Thorne, B., & Henley, N. (1975). Difference and dominance: An overview of language, gender, and society. *Language and sex: Difference and dominance*, 5-42.
- Weatherall, A. (2012). Discursive psychology and feminism. *British Journal of Social Psychology*, 51(3), 463-470.
- West, C., & Zimmerman, D. H. (1983). Small insults: A study of interruptions in cross-sex conversations between unacquainted persons. *Language, gender and society*, 102-117.

APPEDIX
NEWSPAPER ARTICLES

Code	Title of the Article	Author(s)	Newspaper	Mass Shooter
JH01	Aurora movie theater gunman's location finally revealed: A federal prison in Pennsylvania	Mark Berman	The Guardian US Edition	James Holmes
JH02	Judge defends Aurora shooting jurors as victim's mother denounces life sentence	Associated Press in Centennial, Colorado	The Washington Post	James Holmes
NC01	Nikolas Cruz, Florida Shooting Suspect, Described as a 'Troubled Kid'	Matthew Haag and Serge F. Kavoleski	New York Times	Nikolas Cruz
NC02	17 confirmed dead in 'horrific' attack on Florida high school – as it happened	Claire Phipps	The Guardian US Edition	Nikolas Cruz
AL01	Sandy Hook report – shooter Adam Lanza was obsessed with mass murder	Ed Pilkington	The Guardian US Edition	Adam Lanza
AL02	Sandy Hook Elementary shooting leaves 28 dead, law enforcement sources say	Steve Vogel,, Sari Horwitz and David A. Fahrenthold	The Washington Post	Adam Lanza
AA01	Thirteen people dead after gun rampage at Washington naval yard	Dan Roberts, Paul Lewis and Spencer Ackerman	The Guardian US Edition	Aaron Alexis
AA02	Gunman and 12 Victims Killed in Shooting at D.C. Navy Yard	Michael D. Shear and Michael S. Schmidt	New York Times	Aaron Alexis
DPK01	Texas shooting: at least 26 killed at Baptist church in Sutherland Springs	Tom Dart and Ed Pilkington	The Guardian US Edition	Devin Patrick Kelley
DPK02	Gunman Kills at Least 26 in Attack on Rural Texas Church	David Montgomery, Christopher Mele and Manny Fernandez	New York Times	Devin Patrick Kelley

SP01	Mandalay Bay attack: at least 59 killed in deadliest US shooting	Dan Hernandez, Tom McCarthy and Michael McGowan	The Guardian	Stephan Paddock
SP02	Multiple Weapons Found in Las Vegas Gunman's Hotel Room	Ken Belson, Jennifer Medina and John Eligon	New York Times	Stephan Paddock
DP01	Santa Fe shooting: Texas governor confirms 10 people dead and 10 wounded	Tom Dart	The Guardian	Dimitrios Pagoutzis
DP02	In Texas School Shooting, 10 Dead, 10 Hurt and Many Unsurprised	Manny Fernandez, Richard Fausset And Jess Bidgood	New York Times	Dimitrios Pagoutzis
CHM01	Oregon Shooting at Umpqua College Kills 10, Sheriff Says	Dirk Vanderhart, Kirk Johnson and Julie Turkewitz	New York Times	Christopher Harper Mercer
CHM02	Oregon college shooting is all the more reason to carry guns, say local residents	Rorry Carrol	The Guardian	Christopher Harper Mercer
ST01	After a Rampage, Trying to Grasp What Led a Son to Kill	Martin Stolz	The New York Times	Sulejman Talovic
ST02	Scars and Stripes	Ed Vulliamy	The Guardian US Edition	Sulejman Talovic
SHC01	Face of the Campus killer	Ewen MacAskill and Suzanne Goldenberg	The Guardian US Edition	Seung Hui Cho
SHC02	Gunman's brooding disturbed his family	Johnathan Watts	The Guardian US Edition	Seung Hui Cho
JW01	Immigration center gunman sent letter to TV station revealing motive	Anil Dawar	The Guardian US Edition	Jiverly Wong
JW02	Who is Jiverly Voong aka Jiverly Wong? Conflicting picture of Binghamton gunman emerges	Joe Kemp	New York Daily News	Jiverly Wong
RF01	San Bernardino Shooting Kills at Least 14; Two Suspects Are Dead	Adam Nagourney, Ian Lovett and Richard Pérez-peña	The New York Times	Syed Rizwan Farook, and Tashfeen Malik

RF02	FBI: No evidence San Bernardino killers were part of a cell	Doug Stanglin and Kevin Johnson	USA Today	Syed Rizwan Farook, and Tashfeen Malik
OM01	Troubled. Quiet. Macho. Angry. The volatile life of the Orlando shooter	Kevin Sullivan and William Wan	The Washington Post	Omer Mateen
OM02	Always Agitated. Always Mad': Omar Mateen, According to Those Who Knew Him	Dan Barry, Serge f. Kovalski, Alan Blinder and Mujib Mashal	The New York Times	Omer Mateen
MYA01	Tenn. gunman used drugs, struggled with clash of faith	Greg Jaffe , Cari Wade Gervin & Thomas Gibbons-Neff	The Washington Post	Mohammad Youssef Abdulazeez
MYA02	Chattanooga gunman came from a middle-class Muslim family	Craig Whitlock & Carol D. Leonnig	The Washington Post	Mohammad Youssef Abdulazeez
MNH01	Fort Hood Gunman Gave Signals Before His Rampage	James C. McKinley Jr. & James Dao	The New York Times	Maj. Nidal Malik Hasan
MNH02	Major Nidal Malik Hasan: Soldiers' psychiatrist who heard frontline stories	Ewen MacAskill	The Guardian US Edition	Maj. Nidal Malik Hasan
LG01	Oakland college shooting: Gunman sought revenge on school, shot victims execution style	Nancy Dillon	New York Daily	One L. Goh
LG02	Psychiatrist deems 2012 Oakland school shooter incompetent to stand trial	Ronnie Cohen	The Guardian US Edition	One L. Goh

A Corpus-based Study of Metafunction in American Presidential Debates

Humera Faraz¹
Lecturer, Air University, Islamabad

Ghazala Kausar²
National University of Modern Languages, Islamabad

Akhtar Abbas³
Lecturer, Air University, Islamabad

Abstract

This study attempts to carry out an analysis of three American presidential debates between Donald Trump and Hillary Clinton that took place before US elections 2016. The debates have been analyzed by applying Systemic Functional Grammar in order to undergo Critical Discourse Analysis by using corpus based approach. Interpersonal metafunction of Halliday's (2001) SFG (now onwards used for Systemic Functional Grammar) has been used to study and further interpret the implications of choosing lexico-grammatical tools like modal verbs, modal adjuncts and personal pronouns. The findings show that interpersonal metafunction is a part of political discourse intending to gain public favors by influencing opinion, establishing speaker's own assertiveness, drawing public attention towards the issues highlighted by them. In addition, the politicians exploit these metafunctional resources of language to defend their positions also. It is also revealed that Hillary Clinton was found more assertive, determined and committed in the course of debates for making more use of modal auxiliaries and first personal pronoun while Donald Trump was found more judgmental, evaluative and straight forward on the basis of frequent use of modal adjuncts and second personal pronoun.

Keywords: assertiveness, corpus analysis, interpersonal metafunction, modality, presidential debates, systemic functional grammar

1. Introduction

Political discourse is constantly being defined and redefined because of its changing face with the advent of globalization. Actually, the process of political communication seems to revolve around the defining issues of this era (Foster, 2010). In his book on 'Political Communication', he calls this a more complex and expansive concept encapsulating the politics of language. In fact, the main objective of political figures is to achieve their goal, to gain public favor, to make people believe in them and to win the situation as well as votes. How do they gain all this through language and

communicative function their language plays? Foster (2010), expounded this situation in other words, on how political parties plan to organize their internal communication in order to realize their political aims. This research paper aims to explore how language assists them to create this function.

Debates and speeches are two different forms among others in public speaking. These are one of the communicative tools in which the speaker interacts with the people through language in order to interact, build relationships, assert, agree, disagree, acknowledge, commit, promise and gain public favor. How do they accomplish these tasks and functions through language has been a focus of this research. A set of three presidential debates of Donald Trump and Hillary Clinton for US elections 2016 have been analyzed with the perspective of how both the candidates express their interpersonal meanings and employ the use of lexico-grammatical structures in their language use. The framework of systemic Functional Grammar (SFG), an approach to conduct grammatical inquiry of functional properties of language was applied on these debates.

Critical Discourse Analysis (CDA) in general involves language use as social practice (Fairclough, 2001) and tends to explore the relationship between language and ideology (Dijk, 1997; Fairclough, 2001; Wodak, 2001). Critical discourse analysis constitutes describing the text, interpreting text and discourse in terms of their relationship with interaction and context and finally explaining the relationship between interaction and the social context (Fairclough, 2001).It also examines production, distribution and consumption of language which is deeply rooted in its communicative and functional properties of language; the approach conceived by Halliday (1985). He endorsed the concept of lexico- grammatical features of language that help in realizing the metafunctions of any language and their relationship with the context on three levels: ideational, interpersonal and textual. Michael Halliday, the pioneer behind this grammar, has always put an emphasis on semiotics, the code of language, and has distanced himself from approaches that make pragmatics a separate field of linguistics divorced from lexico-grammar (Halliday, 1994).

This research has also used the combination of CDA and corpus linguistics in order to find the modality and modulation as features of interpersonal metafunction. The frequency of these expressions identified through Antconc; a software tool has helped the researcher to conclude that who, between Donald Trump and Hillary Clinton happened to be more interpersonal and how.

1.1 Rationale of the Research

Critical Discourse Analysis in this research has been used differently because of its criticism on providing '*fragmentary and exemplificatory*' text types (Fowler, 1996), lacking in academic rigor due to subjective analysis (Widdowson, 1995) and weak representativeness by selecting data randomly for interpretation and then making generalizations (Stubbs, 1997) by rarely comparing the linguistic features with norms in a language (Orpins, 2005).

In this way, the use of Corpus linguistics together with Systemic Functional grammar will definitely provide an empirical flavor to the linguistic analysis. This research is exploring interactional/ interpersonal metafunction in the language used by Donald Trumps and Hillary's presidential debates. Hence the approach, based on systemic functional lexico-grammatical analysis, to investigate how frequency of modal auxiliaries and concordance hits can help reveal the interpersonal patterns, specifically modality and its degree to reflect more commitment, responsibility, promises and assertiveness. In this way we also can see what Wodak (2001) says that language gains power when powerful people make use of it.

In this era of globalization, technology can be used to assist and help to its best using in the research work. It will provide the means for analysis for the present research also. CL now onwards used for Corpus Linguistics can largely be useful, especially when accompanied by the two approaches of CDA and SFG (Baroni & Marco, 2009). This can be ascribed to the fact that CL, CDA and SFG have a number of common features when it comes to linguistic analysis. That is, they all entail the collection and analysis of naturally occurring language data. Moreover, CDA, SFG, along with CL, make use of quantitative methodologies attempting to provide explanations for the findings that their research produces. Thus, the corpus linguistic approach can provide linguists with large amounts of existing data along with computational tools and procedures that can identify quickly and accurately the features researched. This study will be an addition to the researches done on political discourse particularly recent studies in the genre of speeches and debates.

2. Literature Review

In the course of literature review, some studies will be referred which have used SFG mainly and others which have used corpus linguistics as major approach to analyze data. Although some studies have also been done on other two metafunctions: ideational and textual which would definitely help researcher create niche for conducting her own research in the third metafunction named as interpersonal metafunction. One of the pioneering examples is M.A.K. Halliday's own research study of William Golding's 'The Inheritors' (1973) which illustrated how transitivity patterns associated with the main character in the text revealed his mind make up.

Highlighting the importance of SFG as linguistic theory of 20th century, it must be important to mention the contribution of researchers who played a great role in the development of this theory. Representative works done by the researcher include Martinet (1960) who applied SFG on the interviews from politicians. Leech and Cook (1992) applied on advertising English while Thinbault (1986) worked on ideology and gender reflected from radio broadcast and columns in Australia. Coats (1983) focused her attention on semantic properties of modal auxiliaries and Bybee (1995) checked modality on discourse level. Huang (2001) went through the text of commercial advertisements to explore interpersonal metafunction.

Chilton (2004) tends to employ his view point in the form of question: What does the use of language in contexts we call ‘political’ tell us about humans in general? This fundamental question in the background of this research entails Halliday’s notion of SFL in which he emphasizes upon the socially concerned linguistic framework for examining the links of peculiarities of political thought and behavior with the kind of language use. So, Chilton (2004) situates language practices as a resource through which political speakers ‘imbue their utterances with the evidence, authority and truth’ thereby gaining legitimacy in particular contexts. Language tends to establish the ideology and power through which a political speaker wins his favorite response from the audience (Wang, 2010). He analysed Barak Obama’s presidential speeches in this context. The transitivity and modality taken from Halliday’s SFL, he also concluded that the purpose of these speeches is actually manifested in the use of language practices.

Another study on similar pattern was conducted by Sharififar & Rahimi (2015), in which the object of study was Obama’s and Rouhini’s political speeches at UN in the year of 2013. After undergoing the CDA of those speeches, the authors found how two presidents manifested their power and ideology in the use of language in terms of modality and transitivity patterns.

So, keeping in view the studies mentioned in this literature review, the current study may also contribute in SFG and corpus based approach of doing critical discourse analysis in bringing up the recent combination of SFG and corpus linguistics. It will also add up in the area of modality through which we can see the interactional strategies used by contestants to win the vote and decide who was more interactional.

This research aims to investigate the frequency of modal verbs, mood adjuncts, comment adjuncts and personal pronouns used by Donald Trump and Hillary Clinton in the U.S. presidential debates of 2016 elections. This trend will show the realization of further categories of modulation and modalization of the interpersonal metafunction. Through corpus analysis of these three debates, a discussion will be generated to investigate who was more interactive, assertive and determined etc. between two candidates. Following research questions have been addressed in the current study:

1. How does modulation perform interpersonal metafunctions in presidential speeches of Donald Trump and Hillary Clinton?
2. How does modalization perform interpersonal metafunctions in presidential speeches of Donald Trump and Hillary Clinton?
3. How does modality contribute into political stance of Donald Trump and Hillary Clinton through their presidential speeches?

3. Research Methodology

The data for analysis constitutes three debates which normally are conducted to convince the people to vote the two contestants in the US elections. In fact American presidential debates are held every four years before the elections where the candidates turn up in a hall and share their futuristic plans. The moderators pose questions and ask

them to respond in the given time. Education, services, judgments and integrity, age and experience are compared between the candidates on most significant and controversial issues of the country. All three presidential debates consisting of around 16 thousand words of US presidential elections 2016 are taken to convert into text file in order to run them through Antconc software. The concordance hits were seen over the words/tokens of 5098 in first speech, 4778 in second speech and 6182 in the third speech to find the use of modal verbs, modal adjuncts, comment adjuncts and personal pronouns to evaluate in terms of the greater frequency of these modality expressions.

According to Halliday (1994), modals and modal adjuncts have the potential to carry interpersonal meaning of a clause. Although the ‘Mood’ pattern is also the part of interpersonal meaning negotiated by a clause but the present research is only handling modality analysis.

3.1 Conceptual Framework of Analysis

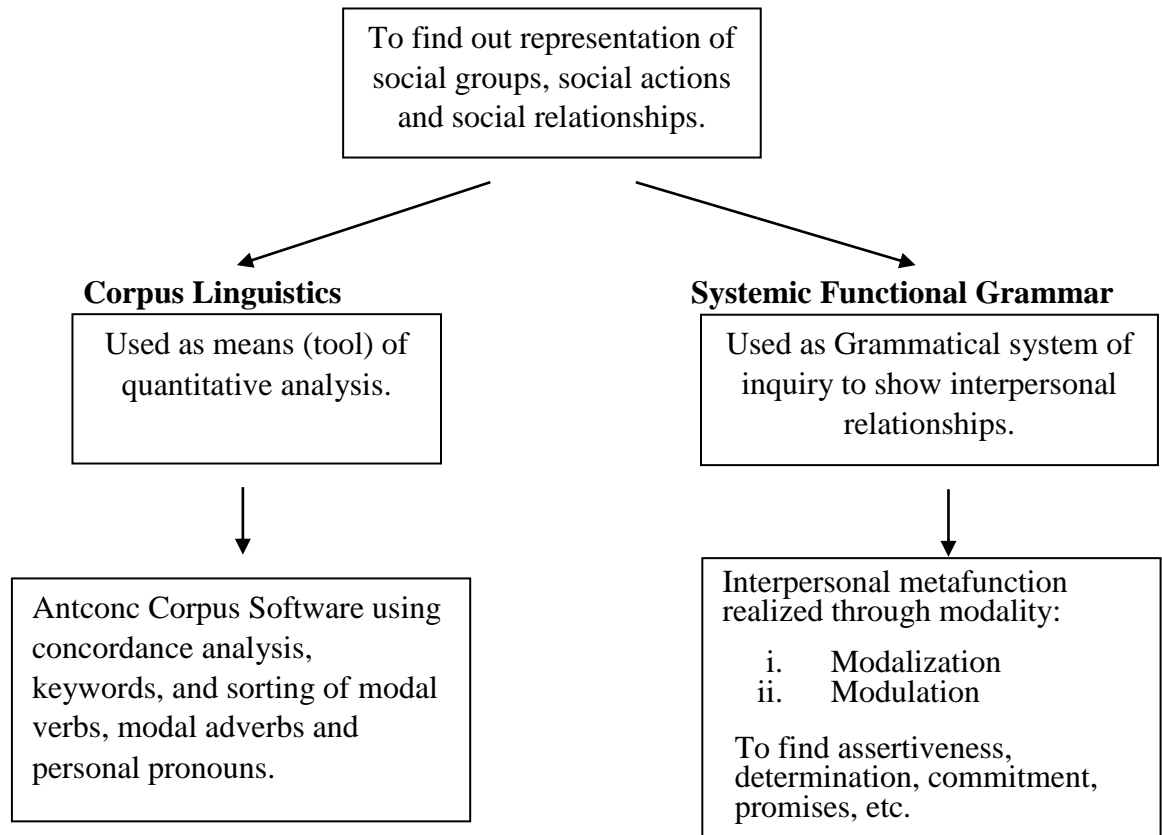


Figure 1: CDA (Critical Discourse Analysis) framework

4. Data Analysis

This section of data analysis presents and demonstrates role of modality in making presidential speeches of Donald Trump and Hillary Clinton interpersonal.

Moreover, the section focuses on displaying the speakers' stance and attitude through modal adjuncts, mood adjuncts, comment adjuncts and person pronouns.

This section focuses on interpersonal role modality including modalization and modulation in political discourse produced by Donald Trump and Hillary Clinton through their presidential speeches.

The presence of a finite in a sentence does not only show time and tense perspective but also the polarity (positive or negative forms) and modality (the intermediate degree between positive or negative forms) for instance, may be might, must, should, could etc. these modal verbs carry interpersonal meaning of modality expressions. This element tends to reflect speaker's emotions, affirmation, futuristic approach, his/her position and attitude towards actions. These modal verbs also reveal speaker's uncertainty, refusal and realization of things and situations. All attributes of modality convey speaker's meaning in the speech (Halliday, 2002a, p. 200).

4.1 Modalization and modulation

According to Thompson (2004), 'Modality' can further be divided into two types, Modalization and Modulation. Modalization shows the judgments of speaker who uses them to validate the propositions which are expressed through probability (possible, probable, and certain) and usuality (usually – something, always). In this process of language tempering, the lexico-grammatical resources are chosen from the modal verbs (can/could, may/might, shall/should, must/ought to and adjective (certain, probable, inevitable, possible, etc.) adverbs (perhaps, rarely, probably, etc.) and other word/phrases (in my opinion, it seems... etc.)

Similarly, the second dimension of modality is modulation which shows the sense of obligation and inclination of the speaker. There is a range of linguistic resources which a speaker chooses to reflect his assertion, confidence and commitment. These linguistic resources range from modal verbs (might/ought to, should, etc.) to adjective (compulsory, mandatory, willing, etc.) adverbs (necessarily, willingly) and other devices (required to, incline, etc.) Mastery of the "situationally appropriate expressions of modality enables the writer or speaker to address the intended audience with skill and exhibit a professional interpersonal competence" (Hyland, 1998, p. 440).

Halliday suggested three degrees of modal commitment: high, medium, and low on the scale. Each degree of modal commitment leads to evaluate different meanings. For instance, the median modals of *will*, *would*, *shall* and *should* have been used by both the candidates, mostly to show assertiveness, obligation, commitment and inclination. In these three debates, the inclination aspect was more dominant in Trump's address carrying 83 hits of the modal verb *will* while Hillary expressed more future centered approach in the discussion of her plans by using 63 times the modal *would*

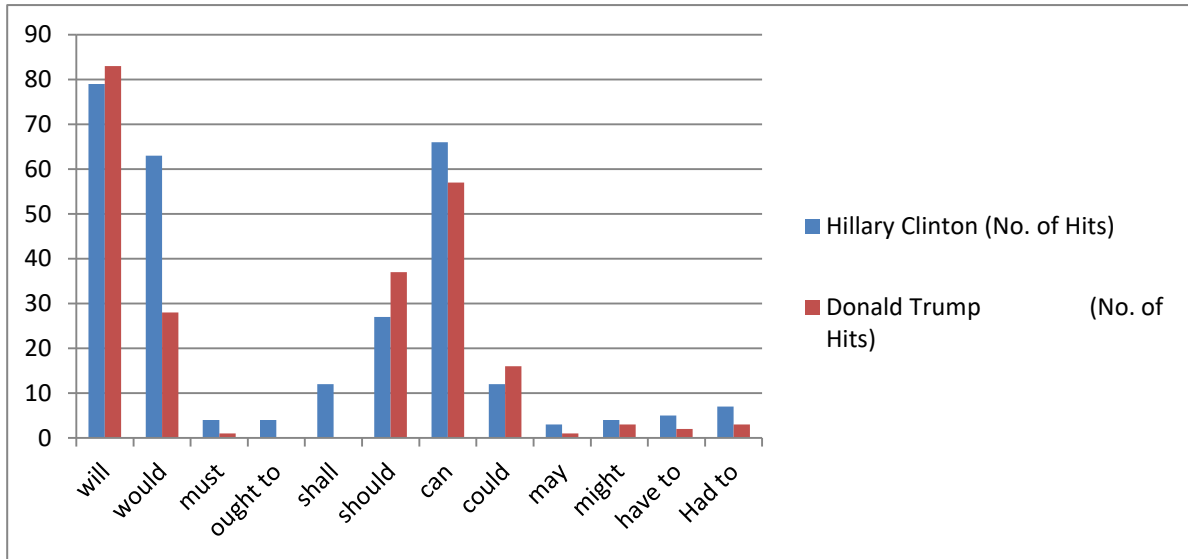


Figure 2: Showing modal verbs used by Hillary Clinton and Donald Trump (comparison)

Donald Trump used finite modal operator (modal auxiliaries) for 225 times to realize different interpersonal meanings while Hillary Clinton used 274 times. Both of the participants used them mostly for conveying the information. A speaker in such kind of debates makes uses of all these scales of modal verbs: high, medium and low. The classification is given below in the table. (Showing Modal auxiliaries and their degree of commitment).

Table1: Modal auxiliaries and their degree of commitment

Polarity	Modality		
	Low	Medium	High
positive	Can, may, could, might	Will, would shall, should	Must, ought to, has to, have to, had to

These different scales of modal auxiliaries show speaker’s attitude that oscillates from high to low level commitments from determination to possibilities and probabilities. It also depicts speaker’s style of showing more or less assertiveness. All of this modality use carries different meanings. Donald Trump used median modal ‘will’ 83 times which is more than the use of ‘will’ used by Hillary Clinton that is 79 times. Similarly, median modals such as ‘would’ have been frequently used by Hillary Clinton (63 times) than Donald Trump (28 times).

‘Shall’ is used by Hillary Clinton (12 times) while Donald Trump did not use even for single time. Now the use of these medians modal operators has been made by both the candidates for different purposes. These modal verbs *would*, *shall*, *should* have been used for marking future tense as well as to illustrate strong wish and desire in accomplishing their works future. It also signaled the certainty of what they said and the

inclination to act what people desire if they are selected. This also shows speaker's assertive yet effective tune to gain public favor and votes.

Low modality auxiliary verbs (*may, might, can, could*) are mostly used by both the candidates to realize the possibilities rather claims, which are the hall mark of politicians to convey this meaning in a polite and effective manner. The low scale modals also show speaker's avoiding personal accountability in case they would not be able to fulfill the promises that they made with people. '*can*' has been used more by Hillary Clinton(66 times) in order to establish the ability with which she can do work for her people if she is elected. On the contrary, Donald Trump used it for 57 times which does not much different from Hillary's use 7 this modal verb. Similarly '*could*' has been used 15 times by Donald Trump which is more than Hillary's use of 12 times. The ratio of the use of the modal verb '*may*' is 3 to 1 and '*might*' 4 to 3 between Hillary Clinton and Donald Trump. Thus, the use of low modal verbs shows speaker's personal beliefs, hopes, possibilities, politeness and likelihood in discourse. For instance:

Clinton says: '*And so what I believe is the more we can do for the middle class, the more we can invest in you, your education, your skills, your future, the better we will be off and the better we'll grow. That's the kind of economy I want us to see again.*'

The use of verbs showing high modal commitment (*must, ought to, has/have/had to*) reflects candidate's firmness, determination and obligation. It also shows their determination to finish the tasks that would benefit people when they are elected. The frequency of use of high scale modal verbs is found more in Hillary's debate than Donald Trump's. Its ratio is as follows:

Must = 4 to 1

Ought to = 4 to 0

Has/have/had to = 5 to 2

Hillary: '*We also have to guide at how we help families balance the responsibilities at home and the responsibilities at businesses.*'

This shows that Hillary is more committed to her claims than Donald Trump. The use of more modal commitments such as *ought to, must* shows Hillary's more authoritative and assertive tone. By using them, she wanted to enable the moderator to show her strong determination on taking actions to attain common objectives.

4.2 Realization of Interpersonal Metafunction through Modal Adjuncts

The modal adjuncts play the roles of expressing degrees of certainty and usuality. They are usually classified into two types, Mood Adjuncts and comment Adjuncts. This graph below shows frequency of modal adjuncts.

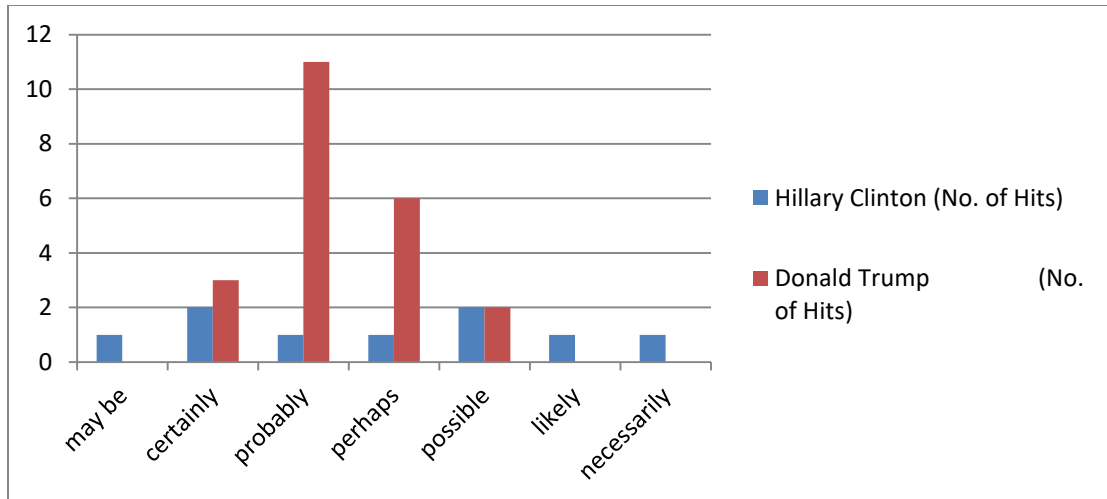


Figure 3: Graph showing comparison in the use of Mood Adjuncts

4.3 Realization of interpersonal Metafunction through Mood Adjuncts

This table and the graph show that overall, Donald Trump used more mood adjuncts especially *probably*, *perhaps* and *certainly* more than Hillary did. As we know that mood adjuncts express the speaker’s attitude and judgments on state of affairs. Both the candidates used number of mood adjuncts to show their own world view. Somewhere in the debate they also lead us to think that the candidates use mood adjuncts to defend themselves if they are pointed out by each other during the debate. For instance,

Trump: ‘and perhaps sadly, Obama probably still listened to you’.
Trump: ‘I agree she was ‘nt there, but possibly she has consulted the peace treaty’

4.4 Realization of Interpersonal Metafunctions through Comment Adjunct

Comment adjuncts are adjectives, adverbs and verbs that show the judgment of speaker’s over state of affairs and actions of him and the other participants.

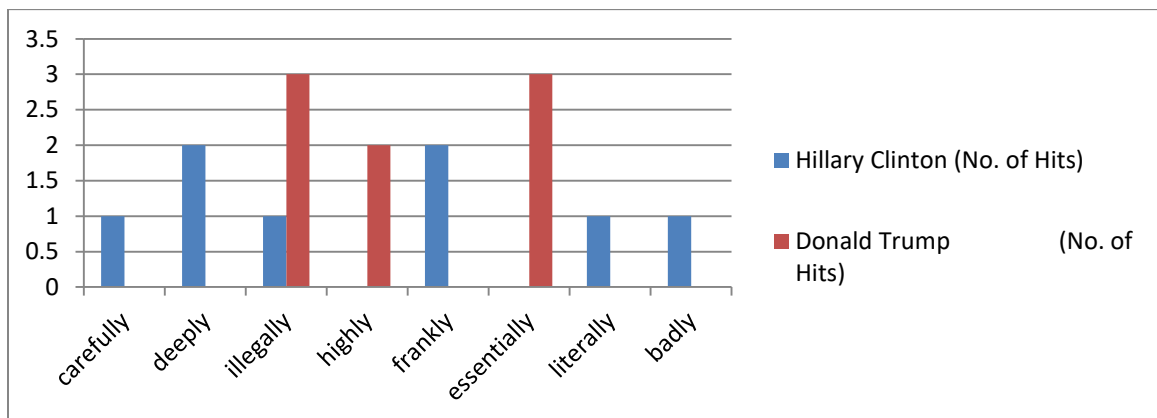


Figure 4: Graph showing comparison in using Comment Adjuncts

Trump: “We have thousands of mothers and fathers and relatives all over the country who are coming illegally”.

Clinton: “.....the Supreme Court and I regret deeply that the senate has not done its job and they have designed to influence our election”.

The use of comment adjuncts helped candidates in convincing people on what they had previously done and what they would do when they are elected. It shows hopeful futuristic approach which has been used equally by Donald Trump than Hillary Clinton.

4.5 Realization of Interpersonal Metafunction through Personal pronouns

Apart from the use modal verbs and modal adjuncts, personal pronouns also carry the interpersonal meaning in discourse. They are used to build a relationship between the speaker and the audience. The use of first personal pronoun *I*, *we* and its analogous *me* and *us* issues different meanings and interpretation as compared to second personal pronoun *you*. According to Mulderrig (2011) and Thompsons (2004) the first personal pronouns *I* and *me* are exclusive which refers to only the speaker and exclude audience while the plural *we* and *us* are inclusive of speaker and audience both. It makes sense of togetherness. The second personal pronoun *you* are exclusive and targets audience.

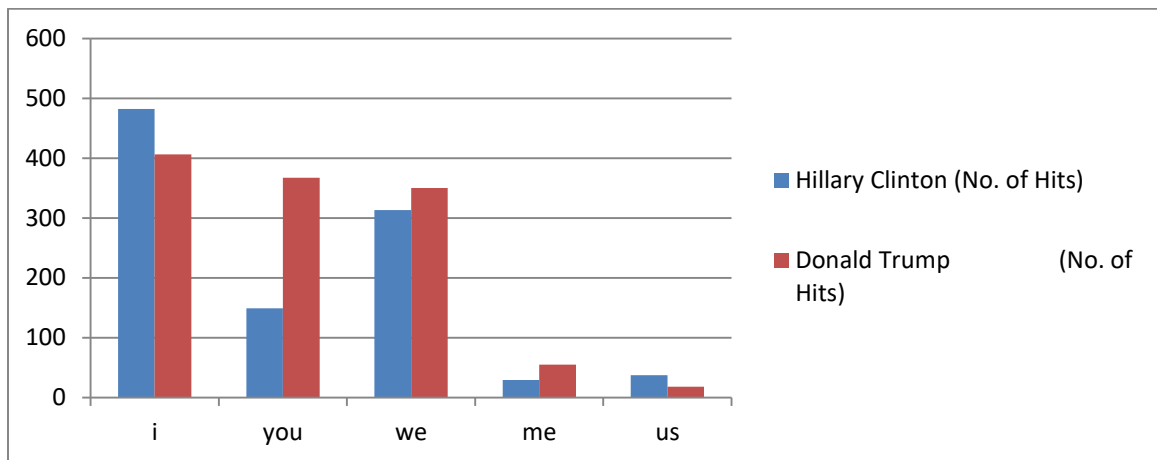


Figure 5: showing comparison in the use of personal pronouns

This graph has revealed that Hillary Clinton use of *I* for 482 times, *me* 29 times, *we* for 313 times and *us* for 37 times. Similarly she used *you* which refer to audience, pointing Donald Trump and US people for 149 times. On the other hand, Donald trump used, *I* for 406 times, *me* for 55 times, *we* for 350 times and *us* for 18 times. He used the second personal pronoun *you* for 367 times which is almost more of the double times than Hillary Clinton used.

Halliday (2000) suggests that personal and possessive pronouns also contains interpersonal meanings owing their ability to build up social relationship between speaker and the audience during a communicative activity. So the choice of personal pronouns by

the speaker especially in a debate session, is very central as it reflects the kind of attitude he wants to stipulate to them as well as to win the favor of audience.

If we take some instances where both the candidate used first personal pronouns and the anamorphous (*I, me*), (*we, us*), it seems that both of them intentionally used the plural form (*me, us*) to arouse the feelings and to create togetherness so that the effect of sharing some objectives can be realized by the US people.

Another reason why both of them used the plural personal pronouns as mentioned below was to shorten the distance between them and the audience. As a matter of fact this was referred to US people so audiences as well as viewers, all were the part of 'we' and 'us'.

Trump: 'What ISIS is doing to us, is a very, very bad thing?'

Hillary: 'I kind of assume, I have a positive and optimistic view of what we can do together.'

From the above examples we can see that the use of *we* and *us* is inclusive of the speaker and the US people. The use of *I* on the other hand implies sense of authority and assertiveness. This refers to power of the speaker to do something with his/her team members. In these three debates, Hillary Clinton used *I* 482 times while trump used 406 times. This shows Hillary's assertiveness and authority which further leads to the representation of white elite American female by her. She also used ' on many occasions when she was explaining her precious experiences as senator.

Hillary: 'I can promise you I will work with every American.'

Hillary: 'I happen to — I happen to support that in a way that will actually work to our benefit. But when I look at what you have proposed, you have what is called now the Trump loophole, because it would so advantage you and the business you do. You've proposed an approach that has a...'

Trump: 'I am a great believer in all forms of energy, but'

In these three debates both the participants used second personal pronoun 'you' to create a dialogic style (Li, 2000) and to separate themselves from each other and the audience.

Donald trump used *you* 367 times which Hillary used for 149 times which is quite less as compared to Trump. Trump most of the times used it to address Hillary and attack Hillary for saying her being unable to become US president. This trend can be related to Donald Trump as if representing white elite American male who target women and may attack their competitors by addressing them directly. This also shows assertiveness but more of an attitude towards others.

Trump: 'why are you (Hillary) thinking about these solutions right now?'

Trump: 'you (Hillary) can't bring back jobs.....'

Hillary also used 'you' for same purpose but somewhere used it to get favor of US people, audience and viewers for making her as their true representative.

Hillary: 'Your child had asthma; you may not be able to have insurance because you can't afford it. So let's fix what's broken about it, but let's not throw it away and give it back to the insurance companies.'

5. Conclusion

This research concludes that the presidential candidates exploit lexico-grammatical structures of modality to win the hearts of people and to convince them to give votes. Throughout three debates, Hillary Clinton and Donald Trump used modal auxiliaries to show strong determination by the help of high modal auxiliaries (*must, ought to, have to, had to*) while using median modal verbs to reflect inclination towards huge beneficial task for Americans and to show commitment of what they were saying will be fulfilled in future. By using low modal auxiliaries, the candidate showed the possibilities as well as abilities to accomplish the most important problems of employment, health leadership, business, trade, immigration issues, governance and foreign relation. While making use of modal adjuncts both the candidates used number of mood adjuncts and comment adjuncts to reflect certainty, likelihood and capability of being evaluative to each other's actions along with their own world views. Use of personal pronouns made both the candidates able to gain public's attention towards their efforts, to shorten the gap between them and the audience/viewers, and to get respect from them to show togetherness. By the use of second personal pronoun, Hillary and Trump both attacked each other on issues of Trump's treatment of womanhood and Hillary's misuse of official emails. From the above discussion, it is concluded that Hillary Clinton seems more determined, obliged, committed and assertive as she has used more modal auxiliaries than Donald Trump who in turn seemed more judgmental, evaluated, attacking and straight forward by using more modal adjuncts. Hillary's debate is inclined more towards 'Modalization' and Trump's towards 'Modulation'. Both of them exploited interpersonal metafunction at different levels to show their status of being American while, elite female versus American white elite male who could be selected as US president of 2016.

References

- Bybee, J., & Fleischman, S. (1995). Modality in Grammar and Discourse. *Revista de Investigacon Linguistics* 2, 131-148. <http://dx.doi.org/10.1075/tsl.32>
- Chilton, P. (2004). *Analysing political discourse: Theory and practice*. New York: Routledge.
- Coates, J. (1983). *The Semantics of the Modal Auxiliaries*. New York, London: Croom Helm.
- Cook, G. (1992). *The Discourse of Advertising*. New York: Routledge.
- Dijk, T. A. V. (1997). *Discourse as social interaction. Discourse Studies*, 2. Los Angeles: SAGE
- Fairclough, N., Mulderrig, J., & Wodak, R. (2011). *Critical discourse analysis*. In *Discourse studies: A multidisciplinary introduction* (pp. 357-378). Los Angeles: SAGE
- Fowler, R. (1996). *Texts and practices. Readings in critical discourse analysis*. New York: Routledge
- Fowler, R., Hodge, R., Kress, G., & Trew, T. (1979). *Language and Control* London: Routledge and Kegan Paul.
- Halliday, M. (1994). *Language as Social Semiotic. Language and literacy in social practice*. England, New York: Multilingual Matters.
- Halliday, M. A. K. (1964). *The Linguistic Sciences and Language Teaching*. London: Longman.
- Halliday, M. A. K. (1973). *Explorations in the functions of language*. London: Edward Arnold.
- Halliday, M. A. K. (1982). Linguistics in teacher education. *Linguistics and the Teacher. Open Journal of Modern Linguistics* 3(1), 10-15.
- Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Hodder Arnold.
- Halliday, M. A. K. (1990). New ways of meaning: A challenge to Applied Linguistics. In *Thirty years of Linguistics evolution*. Amsterdam, Netherland: John Benjamins Publishing Company
- Halliday, M. A. K. (1993). *Language in a changing world*. Australia: Applied Linguistics Association of Australia.
- Halliday, M. A. K., & Martin, J. R. (1981). *Readings in systemic linguistics*: London: Trafalgar Square Publishing.
- Halliday, M. A. K., & Matthiessen, C. (2004). *An Introduction to Functional Grammar*. Beijing: Foreign Language Teaching and Research Press.
- Halliday, M.A.K. (1994). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Huang, G. W. (2001). *Theory and Practice of Discourse Analysis: A Study in Advertising Discourse*. Shanghai: Shanghai Foreign Language Education Press.
- Hyland, K. (1998). *Hedging in scientific research articles* 54. Amsterdam: John Benjamins Publishing.
- Hyland, Ken. Persuasion and context: The pragmatics of academic metadiscourse. *Journal of pragmatics* 30 (4), 437-455.

- Leech, G. N. (1966). *English in advertising: A linguistic study of advertising in Great Britain*. London: Longman.
- Li, L., & Wegerif, R. (2014). What does it mean to teach thinking in China? Challenging and developing notions of 'Confucian education'. *Thinking skills and creativity, 11*, 22-32.
- Martinet, A. (1960). *Éléments de linguistique générale*. Paris: Armand Colin.
- Mulderrig, J. (2011). Manufacturing Consent: a corpus-based critical discourse analysis of New Labour's educational governance. *Educational Philosophy and Theory, 43*(6), 562-573. <http://dx.doi.org/10.1111/j.1469-5812.2010.00723.x>
- Najrzadegan, S., Dabaghi, A., & Eslami-Rasekh, A. (2017). A critical discourse analysis of Iran and US presidential speeches at the UN: The sociopragmatic functions. *Theory and Practice in Language Studies, 7*(9), 764-774.
- Orpin, D. (2005). Corpus linguistics and critical discourse analysis: Examining the ideology of sleaze. *International Journal of Corpus Linguistics, 10*(1), 37-61.
- Sharififar, M., & Rahimi, E. (2015). Critical discourse analysis of political speeches: A case study of Obama's and Rouhani's speeches at UN. *Theory and Practice in Language studies, 5*(2), 343-349.
- Stubbs, M. (1997). Whorf's children: critical comments on CDA. *Evolving models of language 29*, 100-16.
- Thibault, P. J. (1986). Text, Discourse and Context: a social semiotic perspective. *Toronto Semiotic Circle, 3*, 26-42.
- Thompson, G. (2004). *Introducing Functional Grammar*. Beijing: Foreign Language Teaching and Research Press.
- Wang, J. (2010). A critical discourse analysis of Barack Obama's speeches. *Journal of language teaching and research, 1*(3), 254-261.
- Widdowson, H. G. (1995). Discourse analysis: a critical view. *Language and literature, 4*(3), 157-172.
- Wodak, R. (2001). The discourse-historical approach. *Methods of critical discourse analysis, 63-94*.

An Analysis of Writers' Moves and Authorial Stance in Research Articles' Abstracts

Amna Hafeez¹

MPhil Scholar, Air University, Islamabad

Zunera Malik²

MPhil Scholar, Air University, Islamabad

Ahsan Naveed Javeed Cheema³

Research Scholar

National Defense University, Islamabad, Pakistan

Abstract

Research papers' abstracts act as a window for the readers as they can peek through and establish a precise idea about what holds for them. One hundred research papers' abstracts were selected, fifty each from the genres of engineering and social sciences. Moves analysis was carried out in accordance with their linguistic realizations and the authorial stance of the writers. The theoretical framework consisted of Santos's (1996) postulation of move structure in research abstracts, in which he has mentioned five moves structure for writing an abstract. Whereas, Pho's (2008) model was selected for the linguistic realization of those moves and the analysis of the authorial stance. The writers while writing articles' abstracts for the genre of engineering, focus more on the methodology of their research, use minimal attitudinal linguistic items and claim no authority of the research. On the other hand, writers while writing the research papers' abstracts for the genre of social sciences focus more on the introduction of the research, frequently use attitudinal linguistic items and have adapted to claiming authority of their own research. These differences mark the difference of conventions set by the writers of both genres, for writing research papers' abstracts.

Key words: genre analysis, articles' abstracts, moves analysis, authorial stance, attitudinal adverbs, attitudinal adjectives

1. Introduction

The twentieth century has brought with it the shift from focus on product to the process of writing (Gosden, 1995), and the trend continues till today. Research papers are one of the primary sources through which research can be publicized and shared with others, thereby making them a very important part of the research process. The research paper itself can be further divided into sections like introduction, literature review and each section holding its own significance and importance in the overall research paper. Of these parts, an abstract is the first section and consequently the first part of the article that is read or seen by the reader. Moreover, even on platforms with no open access, the

abstracts are still available so that the readers can assess accurately whether or not they really need the research paper. The abstracts, therefore, have the very important function of acting like magnets – so to speak – that will attract the readers towards the actual article (Sandor, 2007; Tseng, 2011). As such, there are certain conventions that need to be followed when penning down the abstract (Donesch-Jezo, 2011), and these conventions are genre-specific; they vary across genres and even across the branches of one genre (Samraj, 2005). Therefore, there is a need to analyze the genre or discipline specific details that an author should follow while penning down a research paper.

The existing research on the structures of abstracts analyzes them from the functional perspective, that is, they look at the different sentences and analyze the function they are performing (Swales, 1990; Santos, 1996; Martin, 2003; Pho, 2008; Tseng, 2011; Çandarlı, 2012; Suntara & Usaha, 2013). For instance, the opening sentence could be explaining what the basic premise of the research is, and thus the function would be something like building a foundation for the following research. An analysis of this kind is termed as moves analysis and the functions are called moves (Swales, 1990; Santos, 1996; Hyland, 2000). Little attention, however, has been paid to the actual linguistic features that exist within these moves (Pho, 2008; Tseng, 2011; Çandarlı, 2012). The linguistic structures and patterns that make a move are rarely explored. Even less likely is the exploration of the authorial stance in the abstract (Pho, 2008). Hence this study analyzes the moves present in abstracts taken from two overarching genres: Social Sciences and Engineering. In addition to this, we have also looked at the linguistic features that exist within these moves and act as indicators as to what function the move is performing. Lastly, we have also looked at how these linguistic features are signaling the authorial stance of the researcher in the abstract.

The world today is a place where publishing has become a mandatory work; it is needed for career advancement which will bring in more prestige and improve one's financial situation which will, ultimately, lead to better living conditions (Çandarlı, 2012). Moreover, even after publishing, the number of times the research paper has been cited in someone else's research goes a long way in establishing one's credibility and reputation in the research community. There is a need, therefore, to ensure that people read your work, and one way to do that is by writing a good abstract. The abstract becomes a kind of force that reels the reader in, and therefore it should be written according to some conventions, and these conventions are genre specific (Samraj, 2005). The exploration of the kind of language, linguistic items and devices, and grammatical patterns that constitute a 'good' abstract is a much-needed task, considering the relative lack of research in this area

This research is significant as it provides future researchers with the genre specific moves structures to be followed when writing an abstract, so as to ensure that their article can attract readership and thereby help enhance their credibility and reputation in their research community. Moreover, our study will help the future researchers to understand the subtle differences the choice of linguistic items can make in academic discourse, and thus aid them in choosing their words carefully and according to the genre in which their writing lies. Lastly, it will enable researchers to understand

whether or not even a small part like an abstract of an article can go a long way in establishing their authorial stance and thereby showing the extent of their own presence in the written discourse.

2. Literature Review

Genre analysis has been the field of interest since the early 20th century and has been developing ever since (Jensen, 2005), its prime focus being on generic constructs and discourse communities, it helps in analyzing the production of different genres within various contexts and their respective interpretations (Bhatia, 2015). All these definitions of genre analysis can be summed up in the idea that genre analysis looks at the linguistic items, features and structures used within different genre writings, and analyzes the functions that are being performed through the use of these devices, while keeping the audience in mind (Bhatia, 2015; Çandarlı, 2012; Jenson, 2005). Under the umbrella of genre studies, Swales' C.A.R.S model (1990) led to the study of 'moves' used by the authors in writing articles' abstracts and introductions. His model paved the path for understanding the certain pattern organized by the writers while introducing their articles (USC Libraries, 2018). Afterwards, research and work were put into the 'moves' analysis of the 'abstracts' of the articles in order to trace the conventional pattern used by the writers within the particular genre.

Research discourses, or analyses of researches themselves, have risen to the top very quickly because of the need for conducting and disseminating quality research, and because publishing is an important phenomenon in the professional world and a requirement for career advancement (Hyland, 2009; Çandarlı, 2012). Under this area of researching the research already done, research articles hold a central position as they are the main platforms where the research is disseminated and publicized (Çandarlı, 2012). Moving further, under the umbrella of research articles, the abstracts are of paramount importance, as it is through the abstract that the researcher will be able to attract the readers and to make them read the whole article and then cite it or use it in their own research (Hyland, 2009). Therefore, the importance of research abstracts within the dimension of the research on research writing cannot be refuted.

An abstract of a research article is the beginning of the 'body' of that article, summarizing it in a paragraph. "The abstract has increasingly become an essential part of a research article" (Pho, 2008, p.231) as it gives the readers an insight about what information the particular article will unfold. It summarizes the article followed, in such a manner so as to make it easier for the reader to make a decision on whether they need to read the article. For instance, they can decide whether the following article is relevant to their interests or whether it focuses on the information they are in search for or not, and if the reader is uncertain about whether the article will be of use to them, the abstract helps them make a decision (Santos, 1996). Santos (1996) also points out that for the readers who have already read the article, the abstract becomes a sort of brief handout using which they can remember the contents of the article without having to go through it again. Lastly, for readers only interested in one part of the research – for instance, the results – the abstract provides a brief description of that part so as to save the readers'

time (Santos, 1996). Other than these practical uses, the importance of abstracts can also be understood by considering the number of journals that publish only abstracts (Rey, 1978), thereby making the abstracts a “highly common and almost obligatory genre” (Santos, 1996, p. 482).

Another vantage point from where the study of abstracts is significant is that the writing of abstracts demands skill and expertise, for it is the abstract which reins in the reader and makes the research article attractive for them (Santos, 1996; Pho, 2008). Therefore, it is a requirement that a kind of convention or a set pattern be formed which the researchers have to abide by when they are writing the abstracts of their research (Harvey & Horsella, 1988). Considering that the rejection rate in some fields and disciplines is as high as seventy-five percent (i.e. 75%), the research abstract is essentially a highly promotional genre which will be the means to ‘sell’ one’s research to the audience, which is why there is a need to examine and analyze this genre and to study the structures of abstracts in different fields (Connor, Nagelhout & Rozycki, 2008). To this end, Harvey and Horsella (1988) have focused their attention towards the formation of a strategic approach to the reading of computational abstracts, where they highlight the structure of abstracts.

Salager-Meyer (1990) has focused on medical English abstracts and proven that even in one discipline, the genre sets vary in the way they structure their abstracts, thereby bringing to light the fact that even within one discipline, abstract writing can have significant differences from genre to genre. Donesch-Jezo (2012) has looked at abstracts from the pedagogical angle and concluded that there are certain rhetorical conventions that the abstracts of a genre follow, and thus abstract writing is something that can be learned and taught if one understands the conventional rhetorical patterns that they are supposed to follow in that specific genre. What these researches show is that there is a certain kind of skill that goes into the writing of a research abstract, and in order to ensure that one’s abstract is doing its job of attracting the readers, there is a need to study the conventions and patterns that are specific to your research genre.

The structure of an abstract can be typically studied by looking at the ‘moves’ it comprises. A move, in an abstract, is a communicative category (Hyland, 2004) that carries a certain kind of information about the article under study; in short, we can say that it fulfils a certain communicative purpose (Hwang, Nguyen, & Su, 2017). For example if a writer refers to the help that would be provided by the article to the readers, then maybe the purpose behind this ‘move’ is to highlight the significance of the study and to indicate the certain ‘gap/loophole’ that the article fills. A move is a ‘discursive’ or ‘rhetorical’ unit that has a particular communicative function within the particular discourse and these communicative functions (moves) are realized with the help of certain steps (Hwang et al., 2017). A move can be considered as a whole unit by itself, as Pho (2008) has done, or it can be divided into several sub-moves or steps, as has been done by Swales (1990) and Santos (1996).

Moves and steps, both, can be obligatory or optional (Hwang et al., 2017). Obligatory moves are more of the conventional or frequently used moves that occur within the abstract relating to a particular genre, whereas optional moves are ones that are

not frequently used in the abstract relating to that particular genre (Hwang et al., 2017). The prevalence in the study of genre analysis has enabled researchers to explore the differences between different academic/non fictional genres and the obligatory moves used by the writers while writing the separate parts of their work. A move analysis, therefore, aims at analyzing the kinds of moves that exist in the writings of a specific genre, thereby bringing to light the conventional move structure of genre-specific writings (Swales, 1990; Santos, 1996; Pho, 2008; Tseng, 2011). Such analysis has been carried out particularly for the ‘abstracts’ of research articles relating to different genres.

Santos (1996) has proposed a five move structure for abstracts, which is what this study will use as the theoretical framework. Using the same five moves structure, Pho (2008) analyzed fifty research abstracts. Pho (2008) extended the moves analysis to include the analysis of the kind of linguistic items and structures used within the moves as well. In this analysis, he explored the function of linguistic features such as verb tense and the modal verbs in identifying the moves that they are being used to represent. In other words, Pho (2008) studied the linguistic features inside each move and identified the typical linguistic markers that indicate certain moves, and penned down the linguistic features that he sees as being most helpful in the identification of moves. Lastly, Pho (2008) also analyzed the authorial stance in the abstracts under scrutiny, examining the move wherein the authorial stance can be most clearly identified. The results of Pho’s (2008) study show that the abstracts of the two genres differ in their moves structure, and that the linguistic items and features used within these moves can be used to distinguish the moves themselves. Lastly, authorial stance is realized in different ways in different moves (Pho, 2008).

In another study, a computational moves analysis of abstracts has been proposed wherein the moves are identified and labeled in light of various rhetorical functions by computer instead of a manual analysis and labelling (Wu, Chang, Liou, & Chang, 2006). Using Swales’ (1990) CARS model, the research puts forward a computational method for the identification of the moves that are there in research abstracts, thereby possibly removing the need for manual analysis. In another work, Suntara and Usaha (2013) have analyzed the subtle differences that exist between abstracts from two related genre systems: linguistics and applied linguistics. Using Hyland’s (2000) model, the study has revealed that there are three conventional moves in the linguistics abstracts while there are four in applied linguistics abstracts. On a similar strain, keeping Pho’s (2008) five move pattern as the theoretical backdrop, Can, Karabacak & Qin (2016) analyzed fifty applied linguistics research abstracts and studied their move structure. The results showed percentage evaluation for the analysis of each category. Their research revealed that the applied linguistics research abstracts omitted the moves of introduction of the topic and discussion of findings. These studies indicate that genre differences do manifest themselves in the abstract – a relatively miniscule part of a whole research article – and thus the analysis of abstracts holds implications – even if they might be small – for the understanding of the overall differences that are there between genres.

The addition of corpus tools in the research area of research on the writing of research has expanded the scope of this area even further and made it easier to handle

large amounts and analyze it. For instance, Ding (2007) conducted a multi-level discourse analysis focusing on a specific genre of application essays written for medical/dental schools. The study was corpus based and conclusions were drawn on the basis of rhetorical objectives of genre. Similarly, Hwang et al. (2017) analyzed scientific research article abstracts, related to nanoscience and nanotechnology, for moves analysis. Hyland's five move model (2000) was used as the theoretical framework, and corpus tools were used for analysis. Similarly, Upton and Cohen (2009) conducted a corpus based study in which they analyzed the rhetorical purpose of selected texts and used the Biber Connor Upton (BCU) Approach based on seven steps for particular discourse analysis. The BCU approach includes the corpus based analysis of the organization of discourse in texts.

The above discussion notifies that corpus based approach to genre study is not new, and has been widely used in the last decade. What's lacking, however, is that the above mentioned studies only analyze the moves and don't study the language within the moves themselves. . The research addresses the following questions:

1. What move structures is being followed in writing of the selected research papers' abstracts?
2. In what ways are the linguistic features within these moves act as indicators for the existence of the moves?
3. How is the authorial stance being reflected through the linguistic features within the moves studied above?

3. Research Methodology

The theoretical framework chosen for our research is the moves structures given by Santos (1996), combined with the linguistic analysis of moves and authorial stance in abstracts given by Pho (2008). A move is considered as a part of a genre which has a certain "minor communicative" purpose which, in turn, fulfills the major communicative purposes of the whole abstract (Santos, 1996). Santos (1996) identifies five main moves in research abstracts: situating the research, presenting the research, describing the methodology, summarizing the results, and discussing the research. He further divides these moves into several sub-moves or steps, and realizes that it is not necessary that all the steps be present in every research abstract. In fact, the differences in the presence or absence of these moves and steps are exactly differentiate different genres from each other. A table explaining the moves' structure, proposed by Santos (1996), is given in table 1.

Table 1: Proposed Pattern for Research Article Abstracts Given by (Santos, 1996, p. 485).

Move 1	Situating the research
	Sub-move 1A – Stating current knowledge and/or Sub-move 1B – Citing previous research and/or Sub-move 1C – Extended previous research and/or Sub-move 2 – Stating a problem
Move 2	Presenting the research
	Sub-move 1A – Indicating main features and/or Sub-move 1B – Indicating main purpose and/or Sub-move 2 – Hypothesis raising
Move 3	Describing the methodology
Move 4	Summarizing the results
Move 5	Discussing the research
	Sub-move 1 – Drawing conclusions and/or Sub-move 2 – Giving recommendations

Using this pattern, Pho (2008) has analyzed the research abstracts from the genres of applied linguistics and educational technology. Along with this generic moves analysis, Pho (2008) proposes a list of linguistic items that carry the potential to act as signifiers or identification markers for the moves they are a part of, as well as indicators of the authorial stance. The list, as given by Pho (2008), is as follows:

Table 2: List of Linguistic Structures to be Analyzed, Given by (Pho, 2008, p. 235-236).

Linguistic structures	Examples/explanations
Grammatical subjects: Phenomenal classes (i.e. what the researcher studies) Class 1: Objects of research and their attributes	‘the participants in the study’, ‘the variables’
Epistemic classes (i.e. nouns “belonging to the researcher or referring to the reasoning of academics” (MacDonald, 1995, p. 158) Class 2: Self-reference Class 3: Other-reference	‘I’, ‘We’, ‘the author’,
• Specific names of other researchers or citations of the author’s own previous studies	‘Swales, 1990’
• Previous research or studies in general	‘previous researches’, ‘research in the area’

<p>without referring to any specific researchers</p> <ul style="list-style-type: none"> • General topics in the field • Specific research objects or outcomes mentioned in previous research 	<p>‘social issues’, ‘mechanics’, ‘automation’ ‘the code was successfully written’</p>
<p>Class 4: Audience</p>	<p>The generic ‘we’ that includes both researcher and audience, ‘one’, ‘you’</p>
<p>Class 5: Reference to writer’s own work – macro-research outcome</p>	<p>‘this study’, ‘this research’, ‘this article’</p>
<p>Class 6: Reference to writer’s own work – micro-research outcome</p>	<p>Referring to the minor sections of the study, such as the research tools, and the like: ‘the findings’, ‘survey’, ‘conclusion’</p>
<p>Class 7: Anticipatory <i>it</i> and existential <i>there</i></p>	
<p>Verb tense and aspect</p>	<p>Past tense; present tense</p>
<p>Voice</p>	<p>Active; passive</p>
<p>Modal auxiliaries and semi-modal verbs</p>	<p>Might, could, may, have to, need to</p>
<p>Epistemic adjectives, adverbs, and nouns</p>	<p>Likely, possibly, certainly, probably, assumption</p>
<p>Attitudinal adjectives, adverbs, and nouns</p>	<p>Important, significant, note-worthy, surprisingly</p>
<p>Self-reference words</p>	<p>I, we, you, our, my, the author (s)</p>
<p>Reporting verbs</p>	<p>Suggest, indicate, point towards</p>
<p><i>That</i>-complement clauses</p>	

This list has been used to analyze the linguistic elements within the moves identified in the abstracts under consideration, to see how much – if any – potential they carry to act as signals for the moves and the authorial stance. However, the list has been delimited to only include the attitudinal adjectives and adverbs, and the authorial stance in our research. These aspects are the ones that can aid us in establishing any existing differences between the two genres selected for study, as the debate between natural sciences’ and social sciences’ authorial stance and opinion-sharing nature is a long standing one. Therefore, we will only be looking at the attitudinal adjectives and adverbs shortlisted from the two corpora under study, as well as the use of ‘I/we’ or ‘the researcher(s)’ that signifies the authorial stance.

Concerning the research design, the study is a qualitative and quantitative content analysis of texts; with corpus acting as a tool for data gathering and compilation and the data being interpreted using qualitative research method. The selected population contains the ‘shortlisted abstracts’ from the genres of social sciences and engineering, published on the Taylor and Francis database, and from this population the sample was chosen using systematic sampling technique. The years 2016 to 2018 was chosen as the time span for the articles. Every second research article was chosen from the results of the chosen time span. The article abstracts were then copied and pasted onto Microsoft Word. After that they were converted into *.txt files and cleaned to remove any features which would act as obstructions for smooth analysis using different corpus software.

The names for the moves to be analyzed in the abstracts were suggested differently in AntMover 1.10 software than the ones suggested by Santos (1996) himself (Table 1) but the functions of the moves were the same. Thus we used moves' names as suggested by Santos (1996) and equalized them with the moves names that appeared in the AntMover 1.10 software on the basis of their functions (Table 3). The moves were analyzed using the software AntMover, which identified the different moves in each abstract. The labels provided by AntMover were then compared with the move structures given by Santos (1996), and were coded according to the labeling given by Santos (1996) by the triangulation method wherein three experts were asked to cross-check the labels we had interpreted for each move. The frequencies of moves were calculated from the articles' abstracts of each genre and chi-square test was run to find the significance of difference between the frequencies of all moves. An alpha of 0.05 was used as the significance's cutoff that is if the probability value was less than 0.05 then it would mean a significant difference in frequencies is present and if the probability value is greater than 0.05 then it would mean that a significant difference in frequencies is not present.

After the identification of each move, the list of linguistic items to be analyzed given by Pho (2008) was used as a guide and those items were analyzed within our chosen texts to see how the linguistic items act as signals for a particular move. In this procedure, the software LancsBox 3.0 was first used to shortlist all the adjectives in the two corpora, from which the attitudinal adjectives were filtered manually. The same procedure was used to shortlist the attitudinal adverbs from the two corpora. Attitudinal adjectives and adverbs are, as the name suggests, adjectives and adverbs that signify or point towards the writer's own personal assessment or opinion; that is, they reflect the writer's attitude towards the subject at hand (Pho, 2008). Items like this were selected from the wordlists of all the adjectives and adverbs present within the corpora. The selected items were then analyzed within their particular context and the differences in the inclusion or exclusion of such items between the genres of Engineering and Social Sciences were analyzed. Lastly, the use of 'I/we', and 'the researcher(s)' was also analyzed in both the genres under study and the usage was compared to establish whether or not any notable differences exist between their use for Social Sciences or for Engineering.

4. Data Analysis and Results

The researchers have analyzed two corpora of fifty research papers' abstracts each, selected from the genres of Social Sciences and Engineering. The researchers have carried out a 'move analyses' by using the corpora and also focused on the linguistic realizations made by the writers in their respective article abstracts. The corpus tools (software) used for achieving goals was 'AntMover 1.10' and 'LancsBox 3.0'.

4.1 Moves Analysis

The terminologies or names used for certain moves, suggested by Santos, differed from the names that the 'ant-mover 1.10' proposed but the communicative acts were the

same. So the researchers equalized the terminologies in accordance with their communicative acts. The following table represents the equalization.

Table 3: Names of Moves Given by AntMover 1.10 and their Equivalent in Santos's (1996) Model.

Moves suggested by the Ant-Mover	Equals (=)	Moves used by Santos
Announcing present research	=	Presenting the research (PTR)
Making topic generalization	=	Situating the research (STR)
Announcing principle findings	=	Summarizing the findings (STF)
Claiming centrality	=	Discussing the methodology (DTM)
Evaluation of research	=	Discussing the research (DTR)

In this study, we shall refer to every move by using the terminologies suggested by Santos (1996).

4.1.1 Patterns of Moves and their Frequency

Table 4 reveals the frequency of different moves used by the authors in the abstracts of individual genres. The abbreviations used to represent the moves have been discussed in table 3.

Table 4: Moves in the Abstracts of the Two Genres under Study and their Respective Frequency.

Genre	Total no of abstracts	PTR (F)	STR (F)	STF (F)	DTM (F)	DTR (F)
Social Sciences	50	50	50	48	46	46
Engineering	50	50	42	44	50	47
Grand Total	100	100	92	92	96	93

Table 4 presents the frequency of moves used by the writers in the article abstracts. It shows that the writers use almost all the moves in the respective articles' abstracts for the readers' better understanding of the content held within the rest of the article. However, the move 'discussing the methodology' (DTM) was used more frequently in abstracts from the genre of Engineering as compared to those from Social Sciences. The reason is that the field of Engineering revolves around technicalities and mechanics that most of readers are not familiar with thus it's important for the writer to thoroughly discuss the methodology used in the research including the tools and the algorithms. Moves' analyses enlightened the researchers that the writers of the abstracts, in genre of Engineering, focused more on discussing the methodology as compared to the rest of the moves. In the field of Engineering, researchers rely on the methods of research so the strength of research depends almost completely on the machines used, the codes written, or the algorithms used. That is why the writers felt it necessary to provide full information about their methodology.

The moves PTR and STR were used frequently in the article abstracts of Social Sciences as the writers focused more on the hypotheses, features and objectives of their research. Thus, the researchers noticed the frequent use of the sub-moves ‘stating the problem’, ‘indicating main features’, ‘hypotheses raising’ and ‘indicating main purpose’ in the Social Sciences research papers’ abstracts. The strength of researches in field of Social Sciences, depends more on the hypotheses and the results and the social work the research can contribute to.

The grand total of the abstracts of both genres, in table 4, describes the total number of abstracts from both genres that is 100. All of the 100 abstracts include the move “Presenting the research (PTR)” which proves that it is one of the most important move observed by the writers while writing the articles’ abstracts. 96 out of 100 abstracts included the move “Discussing the Methodology (DTM)” as methodology binds the whole research together and creates the link between all the chapters in an article thus it is important to be mentioned in the abstract. 93 out of 100 articles’ abstracts observed the move “Discussing the Research (DTR)” and 92 out of 100 abstracts included the move “Summarizing the findings (STF)” as many researchers tend to discuss findings as less as possible so that the reader is compelled to read the whole article out of anticipation. The percentage of moves’ frequency is presented in the table 5:

Table 5: Percentage of Moves’ Frequency.

Moves	PTR	STR	STF	DTM	DTR
Percentage (Social Sciences)	100%	100%	96%	92%	92%
Percentage (Engineering)	100%	84%	88%	100%	94%

Table 5 presents the occurrences of moves and percentages in abstracts. Figure 2 represents the data of table 5 and the occurrence of moves in the articles’ abstracts of both genres parallel to each other. The percentage representation of data further clarifies the discussion above, regarding table 4.

Figure 2, we can see that the move ‘situating the research’ (STR) has been repeated after the move PTR, but as the goal of the researcher has been fulfilled by starting the abstract with which move once, the repetition has not been added to the pattern.

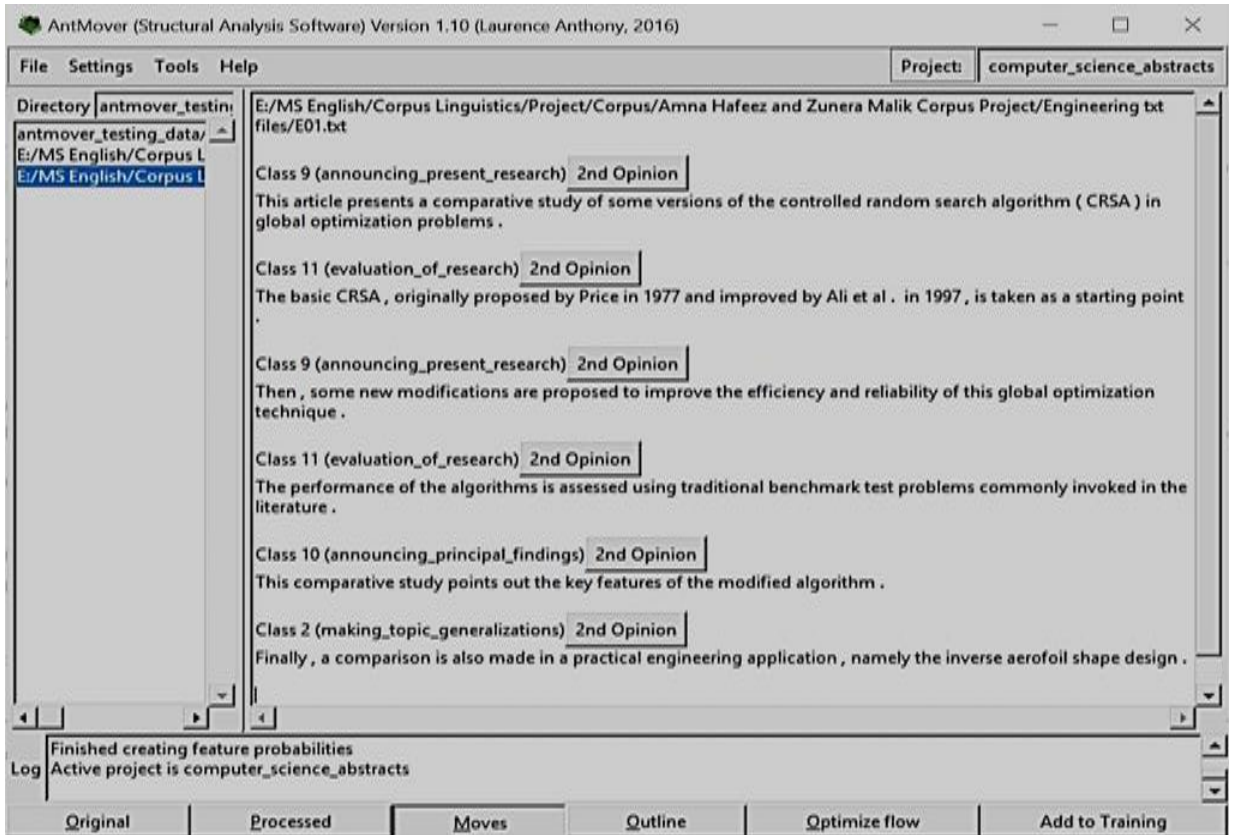


Figure 1: Moves analysis of an engineering abstract as shown by AntMover 1.10.

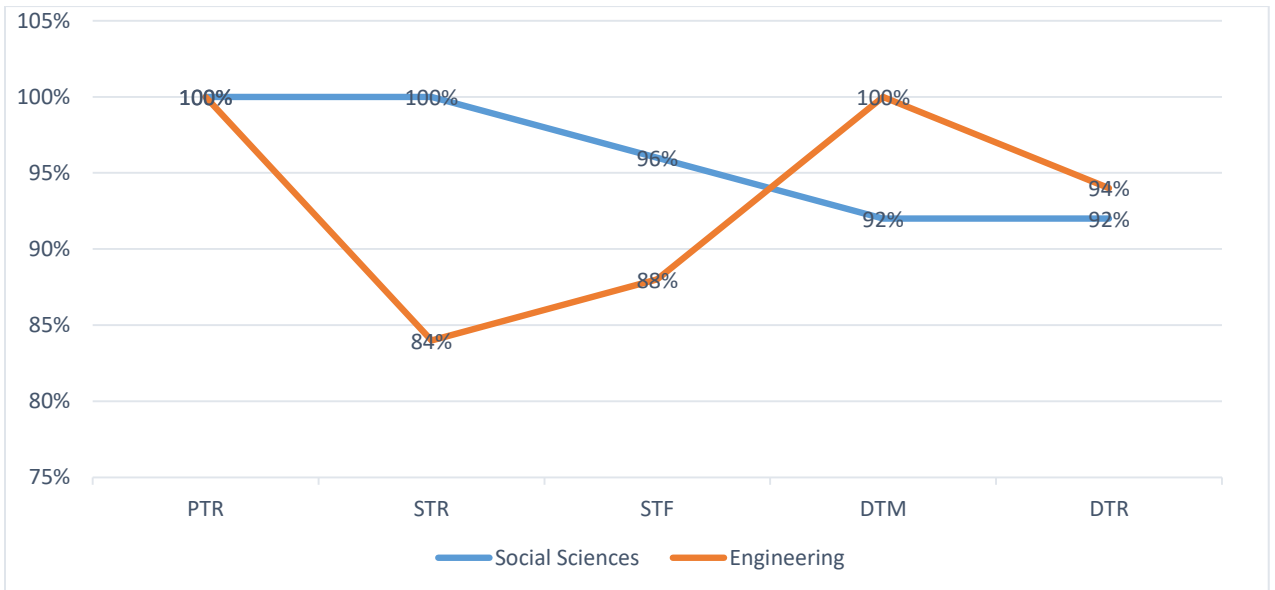


Figure 2: Percentage of Moves' Frequency

The most frequent patterns of moves found in all the article abstracts (both genres) are:

- a. STR-PTR-DTM-STF-DTR
- b. PTR-DTM-STF-DTR
- c. PTR-DTR-DTM-STF

Each pattern represents a unique structure of abstract as chosen by the writer. The pattern (a) includes all of the five moves making the abstract as comprehensive and understandable for the readers as possible. Patterns (b) and (c) do not contain the move STR as the writers preferred to focus more on the purpose and features of their research than to focus on previous literature and its citation or extension.

We found the repletion of moves embedded within the abstracts but the patterns were made analyzing the single usage of moves from start to the end of the abstract.

We applied the chi-square test to the data collected in order to find out whether the difference between the occurrence of moves in the respective genres significant or not. In order to run the test successfully, expected values of moves' frequencies was calculated using the following formula "no of moves occurring*total no of moves in all abstracts/ total no of abstracts". After the calculation of expected values, the chi-square test was run on Microsoft Excel. It was interesting to find out that the probability of difference was greater than 0.05 that is 0.748631987 thus it means that there is no significant difference between the frequencies of moves recorded in the research papers' abstracts of respective genres. So the findings depict that almost all moves are observed and taken into account in the research papers' abstracts of both Social Sciences and Engineering but these moves exist in different patterns (as discussed above) . For different moves a genre based abstract is required.

4.2 Linguistic Realization of Moves

In order to analyze the linguistic realizations made by the writers in the research papers abstracts, we delimited Pho's (2008) framework and analyzed only the authorial stance including self or other reference, attitudinal adverbs and attitudinal adjectives present in the article abstracts.

4.2.1 Attitudinal Adjectives and Adverbs in the Social Sciences Abstracts

Attitudinal adjectives and adverbs were searched using LancsBox 3.0. Using the 'Smart Search' available in LancsBox 3.0, the researchers generated a list of all the adjectives used in the Social Sciences abstracts and Engineering abstracts separately. Out of the generated wordlist, the attitudinal adjectives were picked out manually. This process was repeated for the filtering of attitudinal adverbs. Some of the concordance lines that were picked out as examples of the use of attitudinal adjectives and adverbs are given below. The labels 'SS25.txt', 'SS26.txt' are the names of the individual text files that are present inside the corpus.

SS25.txt: as 2 minutes of practice and practice narratives were particularly beneficial if a unique, rather than commonplace, experience was targeted for

SS26.txt: Tabloid headline descriptors of sexual offenders were also substantially more offensive than those used by broadsheets. In Study 2, tabloid readership
SS26.txt: broadsheets. In Study 2, tabloid readership was associated with more negative attitudes and preferences for harsher punishments for sexual offenders
SS30.txt: self-reported antisocial and psychopathic features in forensic samples, including a significant effect of the assessment instrument and subscale used.

All of these adjectives have been filtered from the wordlist of all the adjectives in the Social Sciences abstracts' corpus. They signify the writer's own attitude towards something; for instance, calling something 'important' or 'significant' is largely a reflection of the writer's own opinion rather than an objective, unrefuted fact. Similarly, words like 'vulnerable', 'traditional', 'negative' are also words which have no definite universal standards or meanings and therefore, are essentially a matter of one's own subjective opinion. Lastly, although they do have a more or less fixed meaning, in the context they are in here, words like 'traditional', 'offensive', 'reasonable' and 'simpler' represent the writer's own attitude towards the subject. For instance, in the concordance line "We focused on how a new and simpler interview strategy, Category Clustering Recall, could increase recall in comparison", the adjective 'simpler' is representing something that is largely subjective and set according to the writer's own opinion.

There is one interesting attitudinal adjective used in the corpus under study: 'dramatic'. There is no doubt that this word is an attitudinal adjective, as dramatic is something which is completely subjective and self-opinionated.

Some concordances, containing attitudinal adverbs, chosen as examples are as follows:

SS16.txt: Ex-intimate stalkers were more likely to be moderately persistent (versus low and high persistence)

SS37.txt: in treatment approach and impact. Effectiveness was primarily defined by reduction in further offending behaviors

SS39.txt: finding showed that the 'assisted' children were significantly more suggestible than the controls with a

SS39.txt: and total suggestibility. In addition, IS was negatively correlated with verbal memory and ability, but

In the case of attitudinal adverbs, the most frequently used adverb is 'significantly', representing the researcher's own views about the subject in question rather than being an objective fact. This information reveals that the Social Sciences' abstracts are more inclined towards subjective, personal opinions rather than plain, objective facts. Consequently, there is more room for personal opinions in Social Sciences than there is in the natural sciences (a part of which is Engineering).

4.2.2 Attitudinal Adjectives and Adverbs in the Abstracts of Engineering Research Papers

The attitudinal adjectives and adverbs in the Engineering abstracts were picked out in the same way they were picked from the Social Sciences abstracts. Some

concordance lines, containing attitudinal adjectives, chosen as examples are as follows. Again, the labels 'E15', 'E16' and the like are names of the individual text files present in the corpus.

E15.txt: automatic train operations, optimal trajectory design is significant to the performance of train operations in

E16.txt: an optimal train-speed trajectory which has equal satisfactory degree on both objectives, a fuzzy linear

E17.txt: It is very important for a manned environmental control system (ECS)

The results indicate that the use of attitudinal adjectives in Engineering research paper abstracts are very low, and what use there is, is there to describe the study as either 'significant' or 'important'. Other than this, the words that occur are 'considerable' and 'useful', which are related mostly to the findings and/or any, changes that happen during/after the experiment. This indicates that there is much less room for personal comments or attitudes to be reflected in the writings.

Some concordances, containing attitudinal adverbs, chosen as examples are as follows:

E24.txt: In particular, the evolutionary game method has potentially faster convergence. This demonstrates the preliminary proof

E28.txt: the maximum fundamental frequency can be increased considerably using variable stiffness design as compared to

E36.txt: and super plasticizer content. The derived model provides sufficiently accurate results for the calibration and verification

The attitudinal adverbs occur even less frequently than attitudinal adjectives. Moreover, the most common adverb is 'significantly' which, like the results for adjectives, relates to the results and/or any changes that occur during the experiment. 'Importantly', 'considerably', 'often', and 'potentially' are some other adverbs that have been used in the Engineering abstracts under study.

Overall, these results indicate that there is lesser room for attitudinal reflections in the abstract of Engineering research papers as compared to the research papers of Social Sciences. The Engineering genre prefers more objective fact sharing than opinionated discourse, while the Social Sciences genre seems more inclined towards the expression of personal opinions.

4.3 Authorial Stance

The researchers analyzed the usage of first person pronouns I/We in the abstracts that indicated the author's claim to his /her authority over the article that is the authorial stance in the abstract. In some abstracts, the writers did not claim their authority and communicated in terms of third person by using phrases like 'the researcher(s)' or 'the writer of this article'.

When it comes to claiming the authorial stance, the researchers have found out that in the field of Social Sciences, 40% (16/40) of the writers claimed their authorial stance in their articles by using terms as 'I', 'we' and 'our'. Examples are as follows:

- i. I argue that seeking to understand...
- ii. Our findings provide no evidence that...
- iii. We also discovered an effect...
- iv. The results support our perceived...

On the other hand, in the field of Engineering, 0% writers claimed their authorial stance in their article abstracts rather they used sentences like:

- i. In this research, an efficient computational framework...
- ii. A gradient-based approach is then established...
- iii. In this article, a hybrid algorithm...
- iv. This article presents a comparative study...

These findings show that the trend of claiming authorial stance has started in the field of Social Sciences but has not yet been adapted by the researchers relating to the field of Engineering.

5. Conclusion

This research has given way to the understanding that researchers from both the academic genres (Social Sciences and Engineering) follow different conventions while writing abstracts for their articles. Although the difference in the frequencies of moves occurring in the abstracts of respective genres is not significant but it is clear from the results that the patterns of those moves differ in both genres. This difference, although small, is still indicating a difference in conventions set by the writers of the respective genres. Researchers, while writing article abstracts for the field of Social Sciences, focus more on the hypotheses and features of their research whereas the researchers from the field of Engineering focus more on the methodology used by them in their research and the results they accomplish. This is also the reason for the low rate of the usage of attitudinal adverbs and adjectives in the article abstracts of the field of Engineering. Thus, the researchers belonging to the field of Social Sciences use frequent attitudinal adjectives and adverbs while describing the features of their research works. Similar is the case with claiming the authorial stance that the researchers belonging to the field of Engineering do not see to but those belonging to Social Sciences have started the trend of claiming their authority and owning their work.

References

- Bhatia, V. K. (2015). Critical Genre Analysis: Theoretical Preliminaries. *Journal of Language and Communication in Business*, 24(54), 9-20.
- Can, S., Karabacak, E., & Qin, J. (2016). Structure of moves in research article abstracts in applied linguistics. *Publications*, 4(3), 23
- Çandarlı, D. (2012). A cross-cultural investigation of English and Turkish research article abstracts in educational sciences. *Studies about Languages*, 20, 12-17.
- Connor, U., Nagelhout, E., & Rozycki, W. V. (eds.), (2008). *Contrastive rhetoric: Reaching to intercultural rhetoric*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Ding, H. (2007). Genre Analysis of personal statements: Analysis of moves in application essays to medical and dental schools. *English for Specific Purposes*, 26(3), 368-392.
- Donesch-Jezo, E. (2012). Corpus-based teaching of research article abstract writing in L2 courses for university students. *International Multidisciplinary e – Journal*, 1(3), 1-10.
- Dos Santos, M. B. (1996). The textual organization of research paper abstracts in applied linguistics. *Text*, 16(4), 481-499.
- Gosden, H. (1995). Success and research article writing and revision: A social-constructionist perspective. *English for Specific Purposes*, 14(1), 37-57.
- Harvey, A., & Horsella, M. (1988). Exploring and exploiting the structure of computational abstracts. *The ESPecialist*, 9, 229-246.
- Hwang, C. J., Nguyen, T., & Su, T. (2017). Move analysis for scientific abstract sections: a study of nanoscience and nanotechnology research article abstracts. *World Transactions on Engineering and Technology Education*, 15(1), 19-22.
- Hyland, K. (2004). *Disciplinary discourses*. Michigan: Pearson Education Limited.
- Hyland, K. (2009). *Academic discourse*. New York: Continuum International Publishing Group.
- Jensen, U. M. (2005). Genre analysis of system description for pallet handling machine. (Unpublished Master's thesis). Aarhus School of Business, Denmark.
- Martin, P. M. (2003). A genre analysis of English and Spanish research paper abstracts in experimental Social Sciences. *English for Specific Purposes*, 22, 25–43.
- Pho, P. D. (2008). Research article abstracts in applied linguistics and educational technology: A study of linguistic realizations of rhetorical structure and authorial stance. *Discourse Studies*, 10(2), 231-250.
- Rey, L. (1978). *Como Redigir Trabalhos Científicos*. São Paulo: E. Blücher.
- Salager-Meyer (1990). Discoursal flaws in Medical English abstracts: A genre analysis per research-and text-type. *Text – Interdisciplinary Journal for the Study of Discourse*, 10(4), 365-384.
- Samraj, B. (2005). An exploration of a genre set: Research article abstracts and introductions in two disciplines. *English for Specific Purposes*, 24(2), 141-156.
- Sándor, Á. (2007). Modeling metadiscourse conveying the author's rhetorical strategy in biomedical research abstracts. *Revue française de linguistique appliquée*, 12(2), 97-108.

- Suntara, W. & Usaha, S. (2013). Research article abstracts in two related disciplines: Rhetorical variation between linguistics and applied linguistics. *English Language Teaching* 6(2), 84-99.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. New York: Cambridge University Press.
- Swales, J. (2001). Disciplinary discourses: Social interactions in academic writing [Review of the book *Disciplinary discourses: Social interactions in academic writing*] *English for Specific Purposes* 20, 495-498. Retrieved from https://www.academia.edu/7567245/Disciplinary_discourses_social_interactions_in_academic_writing
- Tseng, F. (2011). Analyses of move structure and verb tense of research article abstracts in applied linguistics journals. *International Journal of English Linguistics*, 1(2), 27-39.
- Upton, T. A., & Cohen, M. A. (2009). An approach to corpus-based discourse analysis: The move analysis as example. *Discourse Studies*, 11(5), 585-605.

A Corpus-based Genre Analysis of Newspaper Movie Reviews

Arooj Rana¹

MPhil Scholar, Air University, Islamabad

Junaid Khan²

MPhil Scholar, Air University, Islamabad

Muhammad Zafar³

Assistant Professor, Karakoram International University

PhD Scholar, Air University, Islamabad

Wasima Shehzad⁴

Dean Faculty of Social Sciences, Air University, Islamabad

Abstract

In this paper corpus-based study is integrated with genre analysis that examines the discourse of the reviews of movies written in the newspapers by different reviewers. The paper attempts to elaborate how reviews of newspaper operate by investigating them linguistically and inspects different functions of communication performed with a specific discourse. The conceptual framework of this study is based on the theories of Move Analysis (Swales, 1990; Bhatia, 1993) and the notions of Discourse on the Moves by Biber et al. (2007). Reviews written in *The Express Tribune* and *The News* are collected and a special corpus is created to investigate different moves in movie reviews. The analysis of the movie reviews is done by using corpus tools that focuses on the distribution and function of surface linguistic features including move analysis, exploring the function of communication in movie reviews and studying the commonly found words in each move as well frequent collocates present in the constructed corpus. Based on the results, implications of ESP teaching and some recommendations for future research are made.

Keywords: Biber Connor Upton (BCU) approach, coding protocol, collocates, moves analysis, movie reviews

1. Introduction

Cinema is an indispensable part of modern culture (Michel, Roebbers & Schneider, 2007). In modern society, films are an important part of people's daily lives. Family members, colleagues and friends do talk about the topics, which circulate around movies, on daily basis. Discussion about movies is an integral part of communication these days. The film industry is now compared to other art forms as this industry developed at the end of the 19th century. The popularity of movies has grown rapidly and in today's world, one of the most popular ways to spend free time is to watch a good movie (Young, 2000).

In the United States, around 180 million, viewers watch movies and other videos by using internet connections on monthly basis. The mentioned figure represents 85 percent of internet population in the United States. The dividends earned by box office in North America were about 11.38 billion U.S dollars in the year 2016. In the following year, the revenue amounted to 11.07 billion U.S dollars (Portal, 2018). Movies have an entertainment function, but it must be noted that they also have a great impact on society and people's minds and values. This imperious role of movies in our daily lives adds a lot of importance to movie-reviewers (De Jong & Burgers, 2013).

It is quite evident that for many years viewers of movies thoroughly study the reviews of different reviewers, and then decide whether particular movies are worth their time and attention. In the genre of movies, the reviews of professional media evaluators are of immense importance. They mold the minds of the readers whether to watch a particular movie or not. They inculcate either a positive or negative image in the minds of readers by using their expertise and experience on the information accumulated regarding the movies. (Litman, 1983; Eliashberg, & Shugan, 1997; Basuroy, Chatterjee & Ravid, 2003; Desai & Basuroy, 2005). This is not only the case because film critics are normally adept at their art of changing the perceptions of the readers due to their command over the genre, but they see the movies before their release.

In addition, a very strong relation exists in between a constructive, positive reviews and the number of viewers who decide that the worth of the movie can be calculated against a film's payroll. A study conducted by Reinstein and Snyder (2005) says "... an early positive review increases the number of consumers attending a movie in total over its entire run rather than simply shifting consumers from viewing the movie later rather than earlier" (p. 48). The results of this study are consistent with the study conducted by Simmons (cited in Eliashberg & Shugan, 1997) as it states "A third of the audience sees films on the basis of favorable reviews by critics" (p. 70). The results show that the reviews of movies have a direct and far-reaching impact on the minds of the moviegoers.

The study is carried out by using corpus-based genre analysis (Swales, 1990; Bhatia, 1993). The aim of the investigation is to unearth the underlying elements of the genre, which are explored in the paper. In this study, we have employed corpus-based techniques to make analysis of the reviews in the selected newspapers. The theoretical framework for this study is based on the theories of Move Analysis (Swales, 1990; Bhatia, 1993) and notions of discourse on the Move. A special corpus is created by accumulating articles on movie reviews from well-known newspapers, *The Express Tribune* and *The News*. The analysis, based on corpus, lays emphasis on the distribution and function of surface linguistics features including move analysis, exploring the function of communication in movie reviews and to study the words commonly found in each move as well as frequent collocates present in the constructed corpus.

The reviews, written by experts in movie reviews, can be credible source of English for the students of ESP (English for Specific Purposes) and EAP (English for Academic Purposes) from the point of view of teachers. The topics related to movies always attract student attention and interest. They can learn about new themes, new ideas,

latest fashions and muster up extensive knowledge by reading movie reviews in well-known news articles. It would increase the understanding of English, and the usage of words by professional in the movie reviews.

Ever since movies began showing, reviewers have chosen different ways to view and describe the movies they have seen. The reviewers describe or summarize a work in order to help a diverse audience decide whether to see it (Snyder, 2000). In writing the reviews, the reviewers encode different moves, which perform some communicative functions. Thus, this study attempts to explore how movie reviews in newspapers function by investigating it from the perspective of linguistics and to explore how the function of communication is achieved in a journalistic manner or context by looking at the performance of one particular discourse.

In this study corpus-based analysis has been integrated with discourse analysis to examine the realization of movie reviews in newspapers. The study investigates the pattern of movie reviews for the identification of common moves and the structure of the moves used to portray the information to the audience and further investigation of commonly found words and their collocates by compiling a specialized corpus and using UAM Corpus Tool and AntConc software to analyze the texts quantitatively. Moreover, it would help the ESP teachers to assist L2 learners to enhance and improve their mental lexicon and writing skills (Pedagogical Purposes). It will also help the teachers and learners to understand the genre in movie industry. The paper would also exhibit which content words occur frequently in movie reviews article in Express tribune newspaper.

2. Literature Review

With the advent of 19th Century and gaining popularity of film and motion picture, a new medium appeared in the form of movie criticism and got the equal popularity and fame as the former did (Haberski Jr, 2001). Prolific writers have started writing reviews of the movies, since the audience likes it due to the appealing themes and the knowledge it carries. These reviews are read thoroughly and its importance in the realm of film industry could not be neglected. According to Lopate (2006), "It is arguable ... that in the last fifty years more energy, passion, and analytical juice have gone into film criticism than into literary criticism, or probably any other writing about the arts" (p.73). The role of the critic is very interesting because he places emphasis on the elements that are missed out and writes what should have been added and what should not have been in order to make the work more it engaging. According to Kael (1963) a good critic helps people to understand more about the work than they could see for themselves and a bad critic diminish their interest. The film critics hold the position of innovators and opinion leaders, as they are the leading people who view the new films and share their film-consumption experiences with the public (Chakravarty, Liu & Mazumdar, 2010).

Wyatt and Badger (1984) presented a sample of university students with positive, negative, and ambiguous film reviews, and found that the affirmative relationship between the interest and positive opinion in the review. In another study, Eliashberg and

Shugan (1997) examined the correlation between cash sales and analytical power of film reviews. The study found out that the critics have an imperative role in predicting and influencing the movie's commercial box performance.

Another study by Chen and Xie (2005) determines when and how a film should change its marketing strategies according to ratings. The article uses reviews from consumer magazines in addition to Internet sources, although it recognizes that the Internet has reduced the cost to which consumers can obtain information and is therefore the first source of product ratings.

Moreover, another study accompanied by Reinstein and Snyder (2005) shows that a positive movie reviews contributes in the success of a movie. The results of this study shows a resemblance with the survey report published by Jacqueline Simmons (as cited in Eliashberg & Shugan, 1997), which states that "one-third of moviegoers see movies because of favorable reviews by critics" (p. 70). Moreover, the critic is seen as well educated in the arts and high society, usually incorporating one theory or another into his analysis. By utilizing critical theories and prospects, he/she typically analyzes a movie for a specialized audience (Snyder, 2000).

Boatwright et al. (2007) examine the relation between the intrinsic quality of movies and the movie reviews. In the field of movie industry, the research found out the prominent role of the critics on the market performance of the movies. Moreover, in another study an examination is carried out to see the influence of online commentary and reviews by film critics about the rating of films that will be released soon. The study focuses solely on information that consumers can find online, which can influence their assessment of the movie. It is suggested that consumers will see two reviews that are different and that the consumer chooses one review over the other.

One of the most important goals in better understanding of genre analysis is to deeply understand the conventions of that genre (Bhatia, 1993; Hyland, 2009; Swales, 1990). The better understating of this genre conventions will help the users to improve their knowledge. It also includes the genre of justice (Bhatia, 1993, Hafner, 2010; Propen & Schuster, 2010), business genres (De Groot et al., 2006, Rutherford, 2005; Zhu, 2000). Educational genres (Chapmann, 1999; Freedman et al., 1994) and scientific genres (Hyland, 2009; Swales, 1990). The second most important goal of genre analysis includes testing of the text that claims to be a part of textual conventions of a genre (Bhatia, 1993, Swales, 1990). By doing so, if the text resides in the same conventions, it can be referred to as the same genre. This comparison of different texts groups by applying genre analysis have been in use in the field of intercultural communication investigating if the texts belong same generic structure or different from the original genre (De Groot et al., 2006).

While genre analysis, there are many aspects that may be explores while the main part lies in the communicative goal (Bhatia, 1993; Swales, 1990). According to a claim of genre analysis, every text of the same genre portrays the same communication goals goal (Bhatia, 1993, Swales, 1990), and it is also reflected it some other aspects of text of

genre e.g. style, medium and content. Genre analysis also reflects that writers may use some subordinate and rhetoric functions to get the communicative function of the text, the so called “strategies” and “moves” (Swales, 1990). Any move in a text is used for some specific function (Swales, 1981), such as to explain the movie. A strategy is also the way in which a movement is realized such as by summarizing (part of) the plot.

The major aim of the genre analysis is to comprehend the conventions pertaining to text of the genre which are under discussion (Bhatia, 1993; Hayland, 2009; Swales, 1990). If users of a genre know about textual conventions, then this knowledge can help other to have better understanding of such texts (Bhatia, 1993; Swales, 1990; Upton and Cohen, 2009). It is shown that genre analysis is very productive for different genres including the genre of justice (Prospen & Schuster, 2010), genre of science (Hayland, 2009), genre of business (Zhu, 2000; De Groot et al., 2006) and genre of Education (Freedman et al., 1994). The secondary aim of genre analysis is to evaluate the text that demand to be a fragment of a genre of textual conventions. If a particular text conforms to conventions of particular genre, then it could be claimed that the text belongs to that genre. Various groups of texts have been compared by using genre analysis in different areas of intercultural communication to examine whether texts that are having same genre but coming from contrary or other cultures are same or different in the organization of that genre (De Groot et al., 2006).

In genre analysis, different aspects of a genre could be explored. However, the importance of genre is in its function of communication (Bhatia, 1993; Swales, 1990). Genre analysis claims that each text having similar genre has a common communication functions and is seen in other aspects of a genre such as medium, size or content. It is suggested by genre analysis that the writer can employ various functions of rhetoric and subordination to achieve the goals of communication of the text, also known as the "moves" and "strategies" (Swales, 1990). A move is one part of a huge text employed for a particular function (Swales, 1981) for example, to describe the movie. The following questions are formed for the current study:

1. How do different moves, based on Biber Connor Upton (BCU) approach, are encoded in the newspaper movie reviews?
2. How do moves and communicative purposes of text functioned in the movie reviews text of newspaper?
3. What are the frequent recurring collocates in each move of movie reviews of newspapers?

3. Research Methodology

The following section introduces the conceptual framework as well as the methodology adopted for the analysis of data in order to get to the research objectives of the study. This is of an immense importance as the theoretical concepts provide a guideline for the study and the underlying philosophical assumptions and research strategy underpin the selected methodology and data collection techniques.

This study is carried out by using the method of corpus-based genre analysis for investigating the moves, the lexicogrammatical realization of these moves and their communicative functions in the text. To get to these goals, a specialized corpus is constructed. The conceptual framework used in this study is based on the move based analysis framework the concept of which is given by Swales (1990) and Bhatia (1993). Moreover, this study has also adapted Biber et al. (BCU Approach) (2007).

Swales (1990) sheds light on the definition of genre as “A genre comprises a class of communicative events, the members of which share some set of communicative purposes” (p.58). In the given definition, Swales lays emphasis on the significance of the communicative purpose, because the condition is to differentiate one genre from another. In the field of genre, another contribution of Swales is to look at the genre from a perspective of discourse community and connects writers, readers and social contexts. The concept of the genre allows teachers to go beyond the propositional content, forms and processes of writing so that the teachers could help their pupils in their attempts to communicate with the audience well. The contribution of Swales in the field of genre has a terrific effect on academics, writing pedagogy as well as in research-related fields.

Bhatia (1993) who is prominent in genre analysis has contributed to the dynamics of researches of ESP and further contributes to other discourses that are professional. Bhatia carves out the cognitive dimension of genre reconstruction from the genre definition of Swales and balance the definition of genre: “Genre is a recognizable communicative event characterized by a set of communicative purpose(s) identified and mutually understood by the members of the professional or academic community in which it regularly occurs. (Bhatia, 1993, p. 1).

As compared to the definition of Swales, Bhatia (1993) adds more weight to the psychological aspect of the genre, as it can provide not only valuable explanation of the genre but also “relevant, but not linguistic, responses to the Frequently Asked Questions: Why do the members of the discourse community write the way they do?” (p. 19). A linguistic description of various genres is of immense importance, but it can “reveal very little about the true nature of genre and about the way social purposes are accomplished” (Bhatia, p. 18).

The Biber Connor Upton (BCU) approach is specifically applied in the theoretical framework of this paper, as it is the methodology that analyzes texts in terms of the communicative structures that typically comprise texts in a genre: move analysis. Biber et al. (2007) examine the structure and pattern of discourse could be examined by the use of corpus. In order to get to the results corpus based description of the organization of the structure of movie reviews with the communicative functions of the text is taken as a starting point for the investigation. Biber et al. (2007) give seven

analytical moves that must be followed for the analysis. These steps are referred as Biber Connor Upton (BCU) approach. In our study, we have followed Biber Connor Upton approach to develop “The Coding Protocol of Move Type” in movie reviews.

3.1 Research Design

The nature of this research is exploratory. The study has a mixed-method approach as it analyze the data quantitatively as well as qualitatively. In the quantitative analysis of data, a specialized corpus is constructed and then the statistics are analyzed qualitatively to investigate the linguistics features of movie reviews published in Pakistani newspapers. The review articles that we have chosen are from newspaper *Express Tribune* and *The News International*, which are widely read across the country. The reviews are published on weekly basis, and each article comprises 650 to 1,000 words per article. The reviews are selected from October 2017, to April 2018. We have selected 14 articles, and it consists of 9526 words.

We have used both qualitative and quantitative mode of study to analyze our data through the built corpus. For our research study, first, we have done move analysis of our selected data by using BCU approach. The structure and format of each movie review as well as the frequencies of moves encoded in movie reviews is gauged manually. Moreover, each move is analyzed individually by using UAM Corpus tool. Furthermore, the collocation of each move is also investigated in the study in order to realize the lexico-grammatical features of each move.

4. Data Collection and Analysis

Based on the BCU approach (Biber Conner Upton Approach) an analytical framework (Table 1) has been developed to analyse different moves in the movie reviews. Moreover, coding scheme is done through the analytical framework for the analysis of move. In this framework, there are five different moves consisting of several steps. In order to make the results more authentic and objective two post-graduate students, who have training in move analysis, participated in the study. The inter-coder reliability turned out to be 91.3%

Table 1: The Coding Protocol of Move Type in Newspaper Movie Reviews

MOVE 1 Introduction of the Movie
Step 1a: Introducing the title
Step 1b: Introducing the genre of movie
Step 1c: Introducing the director
Step 1d: Introducing the main characters
MOVE 2 Summary of the Movie
Step 2a: Discuss the setting
Step 2b: Refer to movie’s background
Step 2c: Describe the story-line
MOVE 3 Analysis of Movie

- Step 3a: Cinematography
- Step 3b: Theme around which movie circulates
- Step 3c: Contribution of whole cast

MOVE 4 Creative Elements in the Movie

- Step 4a: Focus on the tone of the movie
 - Step 4b: Discuss the music of the movie
 - Step 4c: Describe the special effects in the movie
-

MOVE 5 Opinion Making

- Step 5a: Discuss the strong points of the movie
 - Step 4b: Discuss the shortcomings of the movie
 - Step 4c: Recommendations
-

4.1 Results

In our study we have merged discourse analysis with Corpus-based text analysis to discern the realization of movie reviews in the newspaper. By integrating BCU approach we shed light on the organization of the structure of the movie reviews to investigate common moves and move structures which are employed to convey information to the audience. We have further investigated the linguistic features of these movements by compiling a specialized corpus and by applying the UAM Corpus tool for further analysis. The communicative purpose of a movie reviews is to address the audience if the movie is good or not and whether they should see it and why or why not? Readers expect the feeling of having experienced the movie themselves and they expect the reviewers give an opinion on whether they should go to watch it or not. After having a deep analysis of structure of 14 selected articles, it is evident that the articles are not structured chronologically, and they have slightly different order in respect to introduction, theme, story-line, cinematography, analysis etc. We have also noticed that in contrast to academic writing, the thesis statement of a movie review article is subsidiary, indirect, and not positioned. In addition to it, each review article comes with a heading. We have observed that among the 14 articles, almost 96.6% of the article headings or subheadings demonstrate the name of the reviewed movie.

Table 2: Frequency of Moves from 14 Newspaper Movie Reviews

MOVE 1 Introduction of the Movie	26.6%
Step 1a: Introduce the title	5.9%
Step 1b: Introduce the genre of movie	2.9%
Step 1c: Introduce the director	10.1%
Step 1d: Introduce the main character	7.7%
MOVE 2 Summary of the Movie	17.1%
Step 2a: Discuss the setting	0.5%
Step 2b: Refer to movie's background	4.5%
Step 2c: Describe the story-line	12.1%
MOVE 3 Analysis of Movie	20.1%
Step 3a: Cinematography	8.9%
Step 3b: Theme around which movie circulates	6.5%
Step 3c: Contribution of whole cast	4.7%

MOVE 4 Creative Elements in the Movie	5.7%
Step 4a: Focus on the tone of the movie	1.7%
Step 4b: Discuss the music of the movie	2.9%
Step 4c: Describe the special effects in the movie	1.1%
MOVE 5 Opinion Making	29.6%
Step 5a: Discuss the strong points of the movie	12.5%
Step 4b: Discuss the shortcomings of the movie	14.2%
Step 4c: Recommendations	2.9%

Here it is imperative to identify and count the occurrence of moves in order to investigate how movie reviews are presented in newspapers. Table 2 provides the result of the move analysis of 14 selected articles. In our analysis we observed that Move 5 (29.6%) occurred most frequently, followed by Move 1 (26.6%) and Move 3 (20.1%). The higher percentage of Move 5 explicitly shows that movie-reviews are more inclined towards providing information regarding good and bad aspects of the movies. By highlighting the both of aspects of movies, the movie reviewers provide the information to the audience with which they can gauge their interest. Taking a more-closer look at the steps, the top-high frequency steps is 5b, which shows that movie reviews are more concerned with highlighting the shortcomings of the movies.

4.2 Analysis of Move 1 Using UAM Corpus Tool

It is evident from the statistics that Move 1 comprises 1821 words. Nouns are used significantly as the statistics read that out of 1821 words the figure *Noun* is 507 which makes 27.84% of the entire move. It clearly stands out, whereas it is followed by *Verb* and there are 234 verbs, making 12.85% of Move 1. Also, there are 159 adjectives, making it 8.73% in Move-1.

Project:	Mo1	
Counting:	global	
Unit:	<Mo1/>	
Date:	Sun May 6 03:29:50 2018	
Feature	N	Percent
TOKEN-TYPE2	N=1821	
word	1630	89.51%
punctuation	187	10.27%
TOKEN-TYPE	N=1821	
noun	507	27.84%
verb	234	12.85%
adjective	159	8.73%
pronoun	108	5.93%
determiner	202	11.09%
adverb	95	5.22%
other	68	3.73%
number	20	1.10%
preposition	183	10.05%
conjunction	58	3.19%

Figure 1: Analysis of Move 1 Using UAM Corpus tool

Project: Mo1
Query: <Mo1 feature="nou
Date: Sun May 6 03:25:5'

film	17
director	12
movie	11
character	8
life	6
plaza	5
role	5
insia	5
drama	5
teacher	4
naina	4
time	4
performance	3
mukerji	3
khan	3
story	3
paddington	3
novel	3

Figure 2: Highest Noun Hits in Move 1

The most frequent noun used in the Move-1 is *Film*, since the whole discussion is about films, so it is on the cards that this word would be encountered more and more by the readers and used a lot of times by the writer. Following are the frequently occurring collocates with the node word “film” on the left side are “Chadwick is the obvious star of the film...”, “She is pretty much on her own in this film...”. From the given instances it gets clear that the whole discussion revolves around films that is why the word *Film* is repeated time and again.

Hit	KWC
1	, Chadwick Boseman is the obvious star of the film. Aamir Khan's roles define him in
2	92s pretty much on her own in this film as she is thrown up against a force
3	directed the original Judwaa as well In this film, director Matt Reeves almost completely drops the
4	n escalating sense of danger throughout the whole film, every now and then, Plaza sucker-punches the
5	Plaza, of the REC series, who elevates this film from what could have been just another '91
6	out. Such is definitely the case for the film in question '96 Veronica Similarly, nostalgia is
7	new movie Hichki hit the theatres recently. The film industry's heavy-weights joined hands together
8	studio: giant monsters versus giant robots. The film is an amalgamation of history's volatile
9	in Shakti. Since the events of the previous film, Paddington(MAIN CHARACTER) (voiced with gentle e
10	doubtlessly masterful, and there are parts of the film that certainly work This factor was the influence
11	Wrinkle in Time is a heartless and clumsy film that covers a tussle between the virtuous and
12	just another '91supernatural-thriller' to a film that thrives on its terrifying imagery. Whether i
13	a force of pure evil. She holds the film together with authenticity and the empathy she br
14	immensely gifted actor, he fails to elevate the film with his charisma and snappy one-liners '96
15	he requires to dictate the pace of the film with his directing. Directed by the great Steven

Figure 3: Collocates of Noun *Film* in Move 1

The 2nd most frequent noun in Move-1 is *Director*, as we talked about adjective directorial in the above table that the role of the director in making a movie is indispensable that is why it is no wonder to see the noun *Director* often in the reviews for he directs the movie, the actors and the whole situation.

Hit	KWIC
1	Syed Atif Ali\x92s directorial venturePari The director, Ali, who co-wrote the script with Muhammad
2	moral values that are truly significant in life. Director and producer Advait Chandan did a great job
3	guilty of the aforementioned, Alex Garland\x92s (DIRECTOR) Annihilation takes the proverbial cake with its
4	influence of writer/director Guillermo del Toro (DIRECTOR)at the helm. Likewise, TV director Steven S
5	work This factor was the influence of writer/director Guillermo del Toro (DIRECTOR)at the helm. Likewis
6	the original Judwaa as well In this film, director Matt Reeves almost completely drops the human poi
7	is really challenging to master. But kudos to Director Paco Plaza, of the REC series, who elevates
8	gle an anecdote without lucidity and correlation. Director Paul Kingand his team have put together the
9	all the crafts of filmmaking. As a skilled director, she throughout her movie compelled not only the
10	directing. Directed by the great Steven Spielberg Director Spielberg\x92sworld building is doubtlessly maste
11	del Toro (DIRECTOR)at the helm. Likewise, TV director Steven S DeKnight, whose resume includes Spartacu
12	of seasons of Daredevil, steps up to the director\x92s chair. Directed by Sidharth P Malhotra

Figure 4: Collocates of Noun *Director* in Move 1

4.3 Analysis of Move 2 Using UAM Corpus Tool

There are 3224 words in total in Move-2. The figure of *Noun* is significant from the rest of the part of speeches for there are 910 Nouns, making it 28.23% in Move 2. Noun is followed by Verb, and there are 419 verbs making it 13.00% of Move-2. There are 238 adjectives making it 7.38% of the second- Move.

Project:	Move 2
Counting:	global
Unit:	<Move2/>
Date:	Sun May 6 03:33:29 2018

Feature	N	Percent
TOKEN-TYPE2	N= 3224	
word	2834	87.90%
punctuation	390	12.10%
TOKEN-TYPE	N= 3224	
noun	910	28.23%
verb	419	13.00%
adjective	238	7.38%
pronoun	160	4.96%
determiner	355	11.01%
adverb	148	4.59%
other	115	3.57%
preposition	349	10.83%
number	30	0.93%
conjunction	110	3.41%

Figure 5: Analysis of Move 2 Using UAM Corpus tool

Project: Move 2
Query: <Move2 feature="noun"/>
Date: Sun May 6 03:34:34 2018

film	14
world	13
story	12
father	12
insia	11
war	9
movie	9
life	9
events	8
t`challa	7
characters	6
meg	6
churchill	6
time	5
khan	5
family	5

Figure 6: Highest Noun Hits in Move 2

The most frequent noun used in the Move-2 is *Film*, since the whole discussion is about films, so it is on the cards that this word would be encountered more and more by the readers and used a lot of times by the writer. The instances taken from the moves are “The film offers detail directorial work...”, “The film had been nominated for...” from the given instances it gets clear that the whole discussion circulates around a film that is why the word *Film* is used a lot.

Concordance	Concordance Plot	File View	Clusters/N-Grams	Collocates	Word List	Keyword List
Concordance Hits 14						
Hit	KWIC					
1	such as these are what make a David film and this movie is oozing with such incidents.					
2	the original movie describe the story-line the film follows the story of a teenage girl named					
3	know about. For its flamboyant presentation, the film had been nominated for six Academy Awards and					
4	makes him reconsider the events of the previous film, his feelings towards mankind and whether or not					
5	as possible. Black Panther might be the 18th film in the Marvel Cinematic Universe, but it is					
6	war between the humans and the apes. The film is in facet more about the inner war					
7	, and will be revealed during the movie. The film itself, however, will teach viewers some importan					
8	of a memory of watching Pacific Rim. The film offers detailed directorial work, electrifying pe					
9	winning the prize Ten years after the first film, Pacific Rim: Uprising shows a world in the					
10	, and religious help to end all this. The film started with Pari's substantial character, but					
11	few years after the events of the last film, War for the Planet of the Apes follows					
12	the aftermath of the events of the first film, where the Kaiju (giant monsters/alien invaders f					
13	saved humanity from the Kaiju in the first film \x96 Jake initially seems like the very opposite					
14	t by various world governments. With the original film \x92s characters having either died or disappeared					

Figure 7: Collocates of Noun *Film* in Move 2

Since Move-2 talks about the story line of the movie, so the world is the 2nd frequent noun. The word *World* helps in explaining the plot, and the settings in which the

movie has taken place. The collocates of *world* are “law enforcement by various world governments” “...via the virtual world known as...”, “...to venture into digital world..”. In each case the term *World* is preceded by an adjective thus making clear the situation and the scenario in which it is used.

Concordance	Concordance Plot	File View	Clusters/N-Grams	Collocates	Word List	Keyword List
Concordance Hits 13						
Hit	KWIC					
1	a new form of law-enforcement by various world governments. With the original film's character					
2	the first film, Pacific Rim: Uprising shows a world in the aftermath of the events of the					
3	sheer pleasure of singing and letting the world know about her passion for her talent. Khan					
4	to escape its dystopian reality via the virtual world known as the OASIS. Real-life desolation is					
5	eyes and ears of her people in the world outside. We also have a couple of meaty					
6	became depressed and unsocial since her world renowned astro-physicist father, Dr Alex Murry (p					
7	meets the eye. To the rest of the world, they are a poor central African country, but					
8	ranted zone. Based on real life events of World War II and the United Kingdom's					
9	his decisions and judgments such as, First World War, India policy and backing of Edward VIII					
10	thus left behind to venture into a digital world, where imagination is the only limit. To determin					
11	importantly, starts seeping into the outside world, which slowly and gradually starts eroding. The g					
12	of his kind is viable, even in a world with such few humans					
13	m to arm the oppressed blacks around the world that would prove to be most mentally					

Figure 8: Collocates of Noun *World* in Move 2

4.4 Analysis of Move 3 Using UAM Corpus Tool

There are 1171 words in total in Move-3. The figure of *Noun* is considerably higher than other part of speeches, and it comprises 351 nouns, thus, making it 26.75% in Move 3. It is followed by the figure of Verb and Move-3 comprises 186 verbs contributing 14.18% to Move-3. Also, the percentage of *adjective* is slightly higher this time in comparison with Move-1 and Move-2 which is 9.38%, contributing 123 adjectives to Move-3.

Project:	Move 3
Counting:	global
Unit:	<Move3/>
Date:	Sun May 6 03:38:29 2018

Feature	N	Percent
TOKEN-TYPE2	N=1312	
word	1171	89.25%
punctuation	139	10.59%
TOKEN-TYPE	N=1312	
noun	351	26.75%
verb	186	14.18%
adjective	123	9.38%
pronoun	63	4.80%
determiner	152	11.59%
adverb	75	5.72%
other	45	3.43%
preposition	127	9.68%
conjunction	43	3.28%
number	8	0.61%

Figure 9: Analysis of Move 3 Using UAM Corpus tool

Project: Move 3

Query: <Move3 feature="noun"/>

Date: Sun May 6 03:39:51 2018

movie	11
character	8
film	6
characters	4
world	4
design	3
cast	3
performance	3
wasim	3
khan	3
work	3
wright	2
viewers	2
depth	2
class	2
role	2
love	2

Figure 10: Highest Noun Hits in Move 3

The word *Movie* is frequently used noun in Move-3. Since the whole discussion is about reviews of the movies, its theme, its plot, its characters, its music and about its settings. Some of collocates of *movie* are "...Naina's character is strong thorough out the movie...", "...the movie also touches upon the slum..." or "...the supporting cast makes the movie shine." From the above instances its gets cleared that the noun "Movie" talks about the either the character, the settings or the supporting cast".

Concordance	Concordance Plot	File View	Clusters/N-Grams	Collocates	Word List	Keyword List
Concordance Hits 12						
Hit	KWIC					
1	struggled to find the missing areas in the movie , along with trying to piece together irrelevant c					
2	ding patriarchy and domestic violence, the movie also focuses on women empowerment					
3	ctiveness in the movie is thus reduced. The movie also touches upon the struggle of slum dwellers					
4	enough for it. Although some parts of the movie feel like a fairy tale, and a few					
5	psee Pannustar as the leading ladies in the movie, however, there is not much depth given to					
6	whose prominence and effectiveness in the movie is thus reduced . The movie also touches upon					
7	of the computer-generated imagery in the movie is worth appreciation . The director\x92s real					
8	I feel thoroughly entertained. With a David movie, one expects loads of drama, an action-packed					
9	ance of the supporting cast that makes the movie shine . All of the female characters are strong					
10	nun, that were thrown at them during the movie . The cast is absolutely terrific. And there\x92					
11	is only used as eye-candy in the movie the film is a celebration of the 80s					
12	s character remains strong throughout the movie, which is a good thing, it constantly overpowers					

Figure 11: Collocates of Noun *Movie* in Move 3

In move-3 the word *Character* is the 2nd most frequent noun. This word talks about the pivotal role of the characters and how they are important not only for the enhancing the grace of the movie but also to attract the audience. Some of collocates of character are “...there is one character that stole everyone’s heart...”, “Wasim is the lead character...” “Khan’s character having the persona...” etc.

Hit	KWIC	File
1	92s character(SIDE CHARACTER). Wasim is the lead character and a rookie, and since Khan is a	Move3.txt
2	Wasim\92s shadow, simply to let her character grow His efforts left the moviegoers baffled as	Move3.txt
3	in the limelight. However, despite Khan\92s character having the persona and the depth to become	Move3.txt
4	persona and the depth to become a leading character, he remains in Wasim\92s shadow, simply	Move3.txt
5	ruckus in class. Even though Naina\92s character remains strong throughout the movie, which is a	Move3.txt
6	role that does not overshadow Wasim\92s character(SIDE CHARACTER). Wasim is the lead character and	Move3.txt
7	forementioned powerhouse performers, there is one character that stole everyone\92s thunder and that	Move3.txt
8	does not overshadow Wasim\92s character(SIDE CHARACTER). Wasim is the lead character and a rookie,	Move3.txt

Figure 12: Collocates of Noun *Character* in Move 3

4.5 Analysis of Move 4 Using UAM Corpus Tool

There are 258 words in total in Move-4, it comprises 80 Nouns making it 31.01% of Move-4, 30 verbs contributing 11.63% to Move-4 and there are 32 adjectives, making it 12.40% of Move-4. The percentage of Noun in Move 4 is higher in comparison with the first three moves.

Project:	move4
Counting:	global
Unit:	<Move4/>
Date:	Sun May 6 03:43:11 2018

Feature	N	Percent
TOKEN-TYPE2	N=258	
word	234	90.70%
punctuation	24	9.30%
TOKEN-TYPE	N=258	
noun	80	31.01%
verb	30	11.63%
adjective	32	12.40%
pronoun	9	3.49%
determiner	28	10.85%
adverb	13	5.04%
other	13	5.04%
conjunction	9	3.49%
preposition	18	6.98%
number	2	0.78%

Figure 13: Analysis of Move 4 Using UAM Corpus tool

Project: move4
Query: <Move4 feature="noun"/>
Date: Sun May 6 03:43:58 201:

songs	4
score	3
film	3
hai	2
music	2
track	2
dialogues	2
alteration	1
beat	1
remakes	1
soars	1
performers	1
phira	1
aspect	1
gary	1

Figure 14: Highest Noun Hits in Move 4

The most frequent noun used in Move-4 is the word *Songs*. Most of the lines taken from Move-4 are related to Bollywood that is why here is frequent use of the word *songs* because the reviewer also shed some light on the songs and talk about it. The collocates of the node word “songs” are “Secret Superstar’s song *Main Kaun Hoon is a delightful track*” and “you cannot miss the familiar beat to the songs such as *Chalti Hai Kya 9 Se 12*”.

Hit	KWIC	File
1	the facial alteration of Gary Oldman. The new songs, even though they are a spin on the	Move 4.txt
2	Main Kaun Hoon is a delightful track. Other songs include Meri Pyari Ammi, an emotional and passion	Move 4.txt
3	you cannot miss the familiar beat to the songs such as \x91Chalti Hai Kya 9 Se 12\x92	Move 4.txt
4	\x92ll Miss You and Nachdi Phira are songs that will also touch your soul. From Whishaw\	Move 4.txt

Figure 15: Collocates of Noun *Songs* in Move 4

The 2nd most frequent word in Move-4 is *score*. In Move-4 it is talking about the score of the song (the most frequent word), and the orchestral score and it is assumed cleared from the given examples: “*The music score was sung by Anu Malik...*” and “*orchestral score providing a distinctive experience to the ears*”.

Concordance			Concordance Plot	File View	Clusters/N-Grams	Collocates	Word List	Keyword List
Concordance Hits 3								
Hit	KWIC	File						
1	of Michael Giacchino's immaculate and original score always infuse the film with a feeling of	Move 4.txt						
2	emporary hip-hop, tribal sounds and an orchestral score providing a distinctive experience to the ears.	Move 4.txt						
3	are perfect in their respective parts. The music score was sung by Anu Malik and you cannot	Move 4.txt						

Figure 16: Collocates of Noun *Score* in Move 4

4.6 Analysis OF Move 5 Using UAM Corpus Tool

There are 3005 words in total in Move-5. There are 641 nouns in Move-5 which is 21.33% of last-move. It is very obvious from the percentage that in move-5 the percentage of nouns in this last move is comparatively lower than the first four moves. There are 515 verbs in move-5, contributing 17.14%, whereas, 268 adjectives are there in the last move making it 8.92% of Move-5.

Project:	mmove 5	
Counting:	global	
Unit:	<move5/>	
Date:	Sun May 6 03:47:53 2018	

Feature	N	Percent
TOKEN-TYPE2	N=3005	
word	2715	90.35%
punctuation	287	9.55%
TOKEN-TYPE	N=3005	
noun	641	21.33%
verb	515	17.14%
adjective	268	8.92%
pronoun	176	5.86%
determiner	330	10.98%
adverb	240	7.99%
other	92	3.06%
conjunction	118	3.93%
preposition	316	10.52%
number	22	0.73%

Figure 17: Analysis of Move 5 Using UAM Corpus tool

Project: mmove 5
 Query: <move5 feature="noun"/>
 Date: Sun May 6 03:48:42 2018

film	29
movie	23
plot	9
characters	8
world	7
varun	6
pari	6
life	5
character	5
mind	4
time	4
cinema	4
issues	4
story	4
scenes	3
things	3

Figure 18: Highest Noun Hits in Move 4

The most frequent noun used in the Move-1 is *film*, since the whole discussion is about films, so it is on the cards that this word would be encountered more and more by the readers and used a lot of times by the writer. The collocation of the node word *film* are “... *very scary low-budget film that packs quite...*”, “...*if the film somehow doesn’t spook you...*”, and “...*the film is worth watching...*” From the given instances it gets clear that the whole discussion revolves around a film that is why the word *film* is repeated time and again.



Figure 19: Collocates of Noun *Film* in Move 5

The noun word *Movie* is frequently used noun in Move-5. Since the whole discussion is about reviews of the movies, its theme, its plot, its characters, its music and about its settings. Following are the collocates of the node word *Movie*: “There are much better movies (*Blade Runner, The Arrival...*”, “...*Marvel movie predicting a typical superhero tale...*” and “...*is how the movie manages to subtly and effectively incorporate the Afrocentric issues of racism...*”

Concordance	Concordance Plot	File View	Clusters/N-Grams	Collocates	Word List	Keyword List
Concordance Hits 12						
Hit	KWIC					
1	struggled to find the missing areas in the movie, along with trying to piece together irrelevant c					
2	ding patriarchy and domestic violence, the movie also focuses on women empowerment					
3	ctiveness in the movie is thus reduced. The movie also touches upon the struggle of slum dwellers					
4	enough for it. Although some parts of the movie feel like a fairy tale, and a few					
5	psee Pannustar as the leading ladies in the movie, however, there is not much depth given to					
6	whose prominence and effectiveness in the movie is thus reduced. The movie also touches upon					
7	of the computer-generated imagery in the movie is worth appreciation. The director's real					
8	I feel thoroughly entertained. With a David movie, one expects loads of drama, an action-packed					
9	ance of the supporting cast that makes the movie shine. All of the female characters are strong					
10	nun, that were thrown at them during the movie. The cast is absolutely terrific. And there's					
11	is only used as eye-candy in the movie the film is a celebration of the 80s					
12	s character remains strong throughout the movie, which is a good thing, it constantly overpowers					

Figure 20: Collocates of Noun *Movie* in Move 5

5. Conclusion and Recommendations

In this study, we have explored the different moves, which are encoded in the movie reviews. Our study extended the method of genre analysis to online reviews of movies. For the purpose of study, we developed a coding protocol based on top-down analysis method. Our inter-coder reliability analysis shows that this new encoding protocol is reliable. By extending the method of genre analysis to the genre of online reviews, we therefore offer an addition to the genre analysis. While other genres have been extensively studied with genre analysis (Bhatia, 2004; de Groot et al., 2006; Hyland, 2009; Rutherford, 2005; Zhu, 2000, among others), the analysis of the movie-review genre is relatively new to genre analysis. Moreover, this study research also helps us to focus on the communicative functions performed by those moves and the information presented to readers through them. In addition to it, the highest percentage of Move 5 in Coding Protocol shows that in Pakistani movie reviews the trend of opinion making is highest as compared to others. The second most-occurring move in the list is Move 1 which shows the trend of introducing the movie in the movie-reviews in newspapers. In addition to it, corpus-based linguistic analyses of moves further illuminates the linguistic usage of different movie reviewers. The recurrent nouns and other vocabulary items provides an overview of language of movie reviews. The results of the study could help in developing part of ESP where the teachers could impart knowledge about movie industry to the potential learners about movies and how to write reviews. The course would be designed in such manners that it might include movie reviews from various newspapers, and it could be analyzed by keeping the move structure in mind. The

materials in these reviews could be used for vocabulary, writing, and reading in the field of ESP. The teachers could provide the students with all the word lists such as adjectives, nouns etc. which are most recurring in the reviews. In order to make the corpus more representative and the findings reliable a bigger corpus could be constructed for the analysis of movie reviews. Moreover, comparisons of movie reviews across the boundaries could also be done to know about the difference in movie patterns and linguistic features.

References

- Barrows, C. W., Lattuca, F. P., & Bosselman, R. H. (1989). Influence of restaurant reviews upon consumers. *Hospitality Review*, 7(2), 10.
- Basuroy, S., Chatterjee, S., & Ravid, A. (2003). How critical are critical reviews? The box office effects of film critics, star power, and budgets. *Journal of Marketing* (67), 103-117.
- Bhatia, V. K. (1993). *Analysing genre–language use in professional settings*. London: Applied Linguistics and Language Study Series.
- Biber, D., Connor, U., & Upton, T. A. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure* 28. John Benjamins Publishing.
- Boor, M. (1992). Relationships among ratings of motion pictures by viewers and six professional movie critics. *Psychological Reports*, 70(3), 1011-1021.
- Chakravarty, A., Liu, Y., & Mazumdar, T. (2010). The differential effects of online word-of-mouth and critics' reviews on pre-release movie evaluation. *Journal of Interactive Marketing*, 24(3), 185-197.
- Chen, Y., & Xie, J. (2005). Third-party product review and firm marketing strategy. *Marketing Science*, 24(2), 218-240.
- De Groot, E. B., Korzilius, H., Nickerson, C., & Gerritsen, M. (2006). A corpus analysis of text themes and photographic themes in managerial forewords of Dutch-English and British annual general reports. *IEEE Transactions on Professional Communication*, 49(3), 217-235.
- De Jong, I. K., & Burgers, C. (2013). *Discourse, Context and Media*.
- Eliashberg, J., & Shugan, S. M. (1997). Film critics: Influencers or predictors? *The Journal of Marketing*, 68-78.
- Freedman, A., Adam, C., & Smart, G. (1994). Wearing suits to class: Simulating genres and simulations as genre. *Written communication*, 11(2), 193-226.
- Haberski Jr, R. J. (2015). *It's only a movie! Films and critics in American culture*. University Press of Kentucky.
- Hafner, C. A. (2010). A multi-perspective genre analysis of the barrister's opinion: Writing context, generic structure, and textualization. *Written Communication*, 27(4), 410-441.
- Hyland, K. (2009). *Academic DiscourseContinuum*.
- Kael, P. (1963). Circles and squares. *Film Quarterly*, 16(3), 12-26.
- Litman, B. R. (1983). Predicting success of theatrical movies: An empirical study. *The Journal of Popular Culture*, 16(4), 159-175.
- Lopate, P. (Ed.). (2006). *American movie critics: An anthology from the silents until now*. Library of America.
- Michel, E., Roebbers, C. M., & Schneider, W. (2007). Educational films in the classroom: Increasing the benefit. *Learning and Instruction*, 17(2), 172-183.
- Portal, T. S. (2018, May 5th). *Percentage of internet users who watch online video content on any device as of January 2018*. Retrieved from Statista: <https://www.statista.com/statistics/272835/share-of-internet-users-who-watch-online-videos/>

- Propen, A. D., & Lay Schuster, M. (2010). Understanding genre through the lens of advocacy: The rhetorical work of the victim impact statement. *Written Communication*, 27(1), 3-35.
- Reinstein, D. A., & Snyder, C. M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *The Journal of Industrial Economics*, 53(1), 27-51.
- Snyder-Duch, J. (2000). *The evolution of media reviewing: A content analysis of film and television reviews in "The New York Times" and the New York "Daily News"*, 1967—1997.
- Swales, J. (1990). *Genre analysis*. Cambridge: Cambridge University Press.
- Verdaasdonk, H. (1987). Effects of acquired readership and reviewers' attention on the sales of new literary works. *Poetics* (16), 237-253
- VTzina, R. (1997). One of the best movies of the year! An experiment on the influence of critics' recommendations on consumers. In D. Arnott, D. Doyle, V. Shaw and R. Wensley (Eds.), *Proceedings of the XXVIth annual conference of the European Marketing Academy*, Vol. 3(pp. 1300-1312). Warwick, Great Britain.
- Wyatt, R. O. and Badger, D. P. (1984). How reviews affect interest in and evaluation of films. *Journalism Quarterly*, 61, 874-878.
- Young, S. D. (2000). Movies as equipment for living: A developmental analysis of the importance of film in everyday life. *Critical Studies in Media Communication*, 17(4), 447-468
- Zhu, Y. (2000). Rhetorical moves in Chinese sales genres, 1949 to the present. *The Journal of Business Communication* (1973), 37(2), 156-172.

Comparison of Split Infinitives in Pakistani and British English: A Corpus-based Investigation

Bisma Butt¹

Department of English Language and Literature, University of Lahore, Chenab Campus

Behzad Anwar²

Assistant Professor, Department of English Language and Literature, University of Gujrat

Abstract

This article investigates split infinitives in Pakistani English and British English using Kachru's (1985) concentric circles framework. Beginning with a brief description of split infinitives, the article explores the most common 'breakers' in Pakistani English (outer circle) and their variations in terms of use as compared to British English (inner circle). By using sketch engine, findings reveal that the use of split infinitives in inner circle and outer circle Englishes exhibit similarities and differences. In Pakistani English newspaper corpus, 'immediately' appears to be the most frequently used breaker in split infinitives under 'RB' adverb tag whereas in BNC (Newspaper), the most common breaker is 'really'. The seemingly contradicting data indicates that the split between inner and outer circle Englishes is not same as initially hypothesized by Kachru.

Keywords: Corpus description, Kachru's concentric circles, Pakistani English, split infinitive

1. Introduction

Recently, the construction of split infinitives in English has attracted extensive scholarly attention; many linguists compare different split infinitives in different varieties (Martín & Barranco, 2014; Jang & Choi, 2014). In English language, a split infinitive or cleft infinitive is a grammatical construction in which a word or phrase comes between 'to' and the bare infinitive verb. Usually an adverb or adverbial phrase comes between them.

Decades ago, this construction was considered as incorrect grammatical form but recently many linguists have studied this construction under variation and World Englishes perspective. This article mainly focuses on split infinitival structures by providing the concise description contributing the growing body of literature. From a World Englishes perspective, the focus is on the frequent infinitival splitters that discuss the prestige of splitting. As the English varieties are increased, the current study focuses on British English from inner circle as *native* English showing inadequacy in explaining the uncountable types of variation in split infinitive structures in other less documented Englishes that have become salient in day-to-day discourse.

Initially, inner and outer circles of Englishes had been thought to be dichotomous

(Kachru, 1985), some have argued that the differentiable lines of these Englishes are not entirely definitive (Graddol, 1997). On the other side, there is another issue that has been considered in contact linguistics, whether it has been influenced by universal or structural principle or it might be a cause of other culture of same ecology. Therefore, it is necessary to focus on this variety (Pakistani English) to explain that in multilingual environment, the inner circle is developing its own norm that is according to Kachru's (1985) concentric circles. Moreover, it can be predicted that behind these developing norms, there may be a reason of identity creation subconsciously (Page & Keller, 1985). Speakers' identity can be assumed by split infinitival structures and are related to extent the speakers need to pressurize the identification. In conjunction with the descriptive investigation, the theory of identification introduced by Page and Keller (1985) is considered to have solid evidences of split infinitives in Pakistani English.

2. Literature Review

Split infinitive construction has been criticized by Comly (1803) and considered inappropriate or unconventional linguistic norm. This assumption against split infinitive was investigated in the era of Latin education where prevalent prescription of self-proclaimed grammarians described it '*how it should be used*'. This would have most likely stemmed from the desire of English living up to the standards of Latin. Some Latin rules were integrated in English, for instance, '*amore*' can be best translated to "to love" – where the former is a single word and the latter comprises of two words.

In the Star Trek television serial, the captain has invalidated the use of split infinitives when he quoted, 'to boldly go where no man has gone before, Yule (2010) has explained that captain Kirk could have opted for 'to go boldly' or 'boldly to go' to avoid splitting the infinitive. However, this case has seen to be abandoned to speak. Over the years, in different researches, different terms were coined for this construction e.g. cleft infinitives or spiked adverbs. However, split infinitive is being used more frequently (Martín & Barranco, 2014).

Richards and Schmidt (2010) defined split infinitives as cases where 'the infinitive has been separated from the base of the verb', thus making the sentence awkward to read. Greenbaum, Leech, and Svartvik (1985) have described the constructions where adverbs are found between 'to' and 'verb'. Moreover, they used two terms, pre- marker and post-marker. The term 'pre-marker' is used when adverb comes before infinitive construction and 'post-marker' is used where adverb comes between the infinitive constructions. Examples are as follow:

- (1) I want really to punch him.
- (2) I want to really humiliate him.

Despite having popularity, these grammatical rules at that time began to popularize mainly in the 19th century (Huddleston & Pullum, 2002). In English usage survey of texts in 1980, four British adults were involved in the meeting which last for 45 minutes, and it was acknowledged that split infinitives had been used 19 times and

that the form ‘to actually + verb’ had the highest frequency. There had been acknowledgement that there were 3 instances of ‘to not + verb’ in discussion and it was considered for sample discourse (Quirk et al., 1985). On the other hand, Leech et al. (2009) discovered that this ‘solecism’ has tripled from 12 to 35 from Lancaster-Oslo-Bergen Corpus (LOB) to the Freiburg-LOB Corpus of British English (FLOB) in diachronic corpora studies and doubled from 33 to 85 from the Brown Corpus of Standard American English (Brown) to the Freiburg- Brown corpus of American English (FROWN).

As a matter of fact, Quirk et al. (1985) noted that not using split infinitives in some cases such as in examples (3) and (4) may be ‘awkward and unidiomatic’. The conscious stop of worrying has been made clear in example (5) of splitting the infinitive ‘to and stop’. Similar to Huddleston and Pullum (2002), split infinitives are only acceptable in two cases if it improves clarity or avoid awkwardness (Quirk et al., 1985). Moreover, in some cases like example (5), split infinitives promote clarity and prevent ambiguity.

- (3) Paul has tried consciously to stop worrying about his future.
- (4) Paul has tried to stop consciously worrying about his future.
- (5) Paul has tried to consciously stop worrying about his future.

It makes the use of the terms pre-markers and post-markers to refer to the two variants i.e. pre-verbals and post-verbals , one where the adjunct precedes to, and the one where it follows (see examples 1 and 2, respectively).

When it comes to frequency, Quirk et al. (1985) and Huddleston and Pullum (2002) noted that split infinitives are most common with subjuncts of narrow orientation, especially in cases where the infinitive is gradable, or where the adverbials mark degree. Examples are as follow:

- (6) She wanted to understand the problem.
- (7) She wanted to really understand the problem.
- (8) She wanted to further understand the problem.

Alliance (e.g. ‘ought to at least try’) and expletives (e.g. ‘to fucking close the door’) may also be involved in split infinitives. Moreover, according to Quirk et al. (1985), superlatives or other expressions can also intensify the focus such as the following examples:

- (9) I was the first person to ever live in that house.
- (10) Mother is now reluctant to so much as speak to her.

There were not only adverbials found in between ‘to’ and the verb whereas in rare cases, noun phrases and prepositional phrases can also be observed in between infinitives (Huddleston & Pullum, 2002).

There had been a reputable scholarly attention for major grammatical reference, however this splitting phenomenon is surprisingly non-existent in some similar ones, including corpus based references like (Johansson et al., 1999). The split infinitive has recovered the worldwide attention then, and now made it an interesting phenomenon to study, especially in the World Englishes framework. Split infinitive structures are very common in American English (Escudero, 2011). Moreover, this study reveals that these splitters are frequent in written texts as compared to spoken. Further explanation on AmE was given by Jang and Choi (2014) diachronically by getting data from three corpora that are COCA, Time Magazine Corpus and Corpus of Historical American English. Both of them observed few things that are: intensifiers are thought to be most common adverb splitters, limited use of adverb in split infinitival constructions and also the extreme use of split infinitives in the year 2000. Same as that of above mentioned study regarding inner circled English, Mitrasca (2009) put focus on studying Great Britain English (GBE) by using COCA, British National Corpus (100 million words, 1980–1993), Brown’s Academic List (one million words, 1960s) and also internet search through google. The study concluded that in 1960s, the usage of split infinitive was low, and also that split infinitives were three times more specified in GBE than AmE. It was also discovered by Mitrasca (2009) that in AmE, double adverbs (e.g. ‘to at least eat’, ‘to not only go’, etc.) in GBE were less common than AmE. Furthermore, in AmE, compound splitters (e.g. ‘to more than double’, ‘to more fully understand’) were also apparently present.

Accepting the aspect of outer circle Englishes towards split infinitives, Martín and Barranco (2014) examined split infinitives in four Asian Englishes i.e. Philippine English (PhilE), Indian English (IndE), Singapore English (SingE) and Hong Kong English (HKE). GBE was the ending point of their study. They made comparison between distribution across written text and spoken in terms of split infinitive constructions as well as discovering the phenomenon according to the view point of variationists. The research questions to be answered for this study are as follows:

1. What are the split infinitives of Pakistani English in newspaper corpus?
2. What are the variant split infinitives used in Pakistani and British English newspaper corpora?

3. Research Methodology

The current study has a scope to deal with Pakistani English (outer circle English) and British English (inner circle English) by using sketch engine. Newspaper corpus is taken for Pakistani English whereas British National Corpus (newspaper section) is used for British English. Both the corpora are comparable. Table 1 demonstrates the selected data with whole number of words for each corpus used for this study. Comparison of corpora using lemma is graphed in appendix, moreover, all proceedings via Sketch engine are taken as screenshots which are displayed in appendix.

Table 1: Description of Corpora

S. No	Corpus Name	Kachru's Circle Info.	Text Type	Time	No. of Newspaper	Total number of Words
1	British National Corpus (BNC) (Newspaper Section)	Inner Circle British English (B.E)	i. Written Books & Periodicals. ii. Written Miscellaneous. iii. Written to be spoken.	July-December 2017	8 different newspapers	41,354,121
2	Newspaper Corpus (N.C)	Outer Circle Pakistani English (Pak.E)	i. Periodicals. ii. Written Miscellaneous. iii. Written to be spoken.	September 2017	7 different newspaper	11, 226, 817

Raw frequencies based on the size of the corpora are adjusted. To capture all the breakers of infinitive structures in corpora, CQL (corpus query language) feature is used and it requires a special inquiry formula. The present study has used the following formula for this purpose i.e. [tag= "TO"] []{1} [tag= "VV"]. By using this formula, 17 different following tags were captured that were not of the scope of this study. Table 2 elaborates these tags as well as targeted tags for this study with frequencies. Therefore, the current study has used the following two formulas i.e. [tag= "TO"] [tag= "RB"] [tag= "VV"] and [tag= "TO"] [tag= "RBR"] [tag= "VV"].

Abbreviations: (VV= verb, base form), (RBR= adverb, comparative), (RB= adverb).

Table 2: CQL formulas for data extraction

S. No	Formula	Scope	Subsequent Tags for BNC	Frequency	Subsequent Tags for N.C	Frequency
1	[tag= "TO"] [{"1} [tag= "VV"]	x	TO VV VV	3, 773	TO RB VV	1854
			TO RB VV	2780	TO VV VV	558
			TO RBR VV	222	TO RBR VV	329
			TO VB VV	67	TO VB VV	7
			TO NP VV	30	TO VVG VV	5
			TO VVG VV	19	TO VH VV	2
			TO VH VV	9		
			TO VVN VV	Hapax		
			TO SENT			
			VV	Hapax		
			TO RBS VV	Hapax		
			TO NN VV	Hapax		
			TO VV2 VV	Hapax		
			TO VBN VV	Hapax		
			TO TO VV	Hapax		
			TO NNS VV	Hapax		
			TO DT VV	Hapax		
TO : VV	Hapax					
2	[tag= "TO"] [tag= "RB"] [tag= "VV"]	✓	No Further Tagging		No Further Tagging	
3	[tag= "TO"] [tag= "RBR"] [tag= "VV"]	✓	No Further Tagging		No Further Tagging	

The analysis of split infinitives similarly follows that of (Martín & Barranco, 2014). As the research objectives are already mentioned i.e. what are the split infinitives used in Pakistani and British English newspaper corpora. More specifically, the hope is to identify the most common ‘breakers’ in infinitives across Pakistani English in relation to the inner–outer circle framework. Parallel to said above objectives, it is also attempted to justify the emerging patterns demonstrating the attempt to build outer circled identity unconsciously in multilingual context.

4. Results and Discussion

Under ‘RB’ tag, a striking variation can be seen in corpora whereas in ‘RBR’, no variation is observed except the number of hits is higher in N.C (Table 3). RB consists those breakers which have frequency count 10 and for this purpose, 36 breakers are caught in newspaper corpus (Pak. English) whereas 33 breakers are caught in BNC. RB category shows variation in terms of frequency and use of breakers. In Pakistani English newspaper corpus, ‘immediately’ appears to be the most frequently used breaker in split infinitives under ‘RB’ tag whereas in newspaper section of BNC, the most common breaker is ‘really’ (Table 3).

Table 3: Frequent breakers of corpora

Formula	Corpora	Breaker	Frequency	Formula	Breaker	Frequency
1 [tag= "TO"] [tag= "RB"] [tag= "VV"]	British National Corpus (BNC)- Inner Circle	really	85	[tag= "TO"] [tag= "RBR"] [tag= "VV"]	further	153
		even	78		better	48
		simply	69			
		fully	64			
		actually	63			
		just	63			
		successively	33			
		properly	32			
		quickly	32			
		suddenly	28			
		so	26			
		completely	22			
		either	20			
		tally	20			
		seriously	19			
		significantly	18			
		slowly	18			
		then	18			
		finally	16			
		temporarily	16			
		somehow	15			
		greatly	14			
		regularly	14			
		publicly	13			

2

Newspaper Corpus (N.C)-Outer Circle

radically	13		
partially	12		
simultaneously	12		
truly	12		
systematically	11		
jointly	10		
personally	10		
specifically	10		
still	10		
immediately	153	further	271
fully	98	better	51
not	67		
successfully	48		
even	46		
personally	43		
first	40		
really	39		
properly	38		
jointly	35		
also	33		
finally	31		
formally	30		
either	27		
quickly	27		
effectively	26		
strictly	25		
only	22		
just	21		
seriously	21		
publicly	19		
eventually	17		
voluntarily	17		
independently	16		
regularly	16		
safely	15		
thoroughly	14		
illegally	13		
completely	12		
gradually	12		
urgently	12		
clearly	11		
officially	11		
simply	11		
always	10		
significantly	10		

Interestingly, the non-traditional adverb breaker that most frequently emerged is ‘immediately’ in Pakistani corpus which is on top most level whereas this adverb is not even in the list of RB in BNC (Newspaper section). Some other cases of this sort are as follow:

- i. not
- ii. successfully
- iii. formally
- iv. also

According to Martín and Barranco (2014) study, the breaker ‘like’, which appeared at the top, was considered a discourse marker as in the following example:

(11) He was like really winded but he managed to like grab her.

<ICE-IRE: S1A-044 #007:1:A>

Surprisingly, ‘like’ breaker is not even present in both corpora in top frequency list, either in BNC, it is present as hapax whereas in newspaper corpus (Pak. English), no single hit is found. Example is given in the following from BNC.

"Written books and periodicals", "all young ones where I was. The older ones used", "to like bully", "the young ones if you were soft and that, and like"

Adverbs ‘RB’ and comparative adverbs ‘RBR’ seem to be dominated in the breakers of Pakistani and British English. The negator ‘not’, as Fitzmaurice (2002) suggests, is not actually a breaker but the negative counterpart of the infinite ‘to + verb’ – that is, ‘to + not + verb’, for this reason, it is not included in this study.

What has been discussed up to this point is done generally, however, it would benefit us to compare the split infinitive usage between inner and outer circle Englishes to offer a better means for comparison of native and non-native English speakers. Table 3 presents the most popular breakers across the inner circle English. Of the top five in British English newspaper corpus (e.g. really, even, simply, fully and actually) only, ‘fully’ and ‘even’ are present in newspaper corpus but on different frequency level. Based on the ranking of the breakers in relation to use, it appears that the most popular breaker in the inner circle is ‘really’ with a frequency of 85, followed by ‘even’ having a frequency of 78, ‘simply’ with 69, ‘fully’ with 64 and ‘actually’ with 63.

In contrast, the outer circle English (Pak. English) reveals a different situation, where ‘immediately’ appears to garner first place having 153 hits while this breaker has only 13 hits in BNC newspaper corpus. The following table shows the usage of adverb ‘immediately’ as breaker in newspaper corpus (Pak. English) and BNC newspaper corpus with top 8 frequent hits.

Table 4: Comparison of Adverb ‘immediately’ as a breaker in Corpora

S. No	‘Immediately’ in BNC newspaper corpus	Frequency	‘Immediately’ in Pakistani English newspaper corpus	Frequency
1	To immediately read	1	To immediately stop	6
2	To immediately join	1	To immediately resolve	6
3	To immediately impress	1	To immediately remove	6
4	To immediately follow	1	To immediately release	6
5	To immediately fly	1	To immediately withdraw	5
6	To immediately divulge	1	To immediately take	5
7	To immediately defend	1	To immediately provide	5
8	To immediately cancel	1	To immediately arrest	5

Table 4 also reveals the use of adverb ‘*immediately*’ in different contexts in outer and inner circle. However, both corpora show that this breaker is negative. Here the question is raised in which context ‘*immediately*’ is used as split infinitive in BNC newspaper corpus. To check the contextual variability, the researcher used dictionary. BNC and Pakistani English corpora have different thesaurus words for ‘immediately’ and the following table shows most frequent thesaurus words and same words for both corpora i.e. (soon, already, first, then).

Table 5: Thesaurus for comparable corpora

Thesaurus for Comparable Corpora			
S. No	Newspaper Corpus	BNC	Similar Words
1	Soon	Also	Soon
2	Finally	Already	Already
3	Already	Now	First
4	Yet	Just	Then
5	Again	Actually	
6	First	Then	
7	Never	Always	
8	Then	Only	
9	Formally	First	
10	There	Soon	

This result shows that ‘*immediately*’ is used in different contexts in both corpora. Moreover, the following words are found as more frequent in thesaurus for targeted words: also, already, now, actually, then, always, just, quickly, suddenly. When these frequent words are cross checked with above table where the usage of ‘*immediately*’ in newspaper corpus is given, it is assumed that last three words in thesaurus (just, quickly, suddenly) may be used as alternative in Pakistani English newspaper corpus. For this purpose, CQL option having formula [tag= "TO"] [lemma = "quickly"] [lemma = "stop"] in sketch engine is used whereas for RB (just, quickly, suddenly), all frequent targeted verbs (stop, resolve, remove, release, withdraw, take,

provide, arrest) came with ‘*immediately*’ in newspaper corpus are alternatively used. The following table of BNC (Newspaper section) reveals the result having 6 frequent breakers (just, either, really, officially, eventually, actually) with ‘stop’ and ‘take’

Table 6: Frequent breakers of BNC modified by most frequent verbs (‘stop’ and ‘take’) of split infinitive structures of Pak. English

S. No	Formula	Frequency
1	[tag= "TO"] [lemma = "quickly"] [lemma ="stop"]	2
2	[tag= "TO"] [lemma = "quickly"] [lemma ="take"]	2
3	[tag= "TO"] [lemma = "just"] [lemma ="stop"]	5
4	[tag= "TO"] [lemma = "just"] [lemma ="take"]	6
5	[tag= "TO"] [lemma = "suddenly"] [lemma ="stop"]	2
6	[tag= "TO"] [lemma = "suddenly"] [lemma ="take"]	4

Manual contextual reading reveals that out of 6 frequent breakers, only ‘just’ is used as alternative option of ‘*immediately*’. Moreover, the frequency of ‘*just*’ in thesaurus checked via sketch engine is 125523. For more authentications, ‘sketch difference’ option is used and the following table reveals the results that ‘*just*’ is modified by following VV (return, begin, start, take, put, stop, remove). Examples are given below:

❖	<i></p><p> 'What, now?' </p><p> 'We've come too far</i>	to just stop	<i>, haven't we?' </p><p></i>
❖	<i>facing you, on to the needles from which you have</i>	to just remove	<i>the sleeve. Replace the sleeve stitches from</i>
❖	<i>to say 34. Sorry. Don't worry. I forgot to ask you</i>	to just take	<i>the handset off. Can you call her again, no just</i>

Figure 1: Frequency of ‘just’ through sketch engine

Table 7: Alternative breakers in BNC with VV modification

S. No	Alternative Breakers in BNC	VV Modification
1	Just	Return, Begin, Start, Take , Put, Stop, Remove
2	Either	Take, Provide
3	Eventually	Return, Agree, Replace
4	Actually	Start, Begin, Take
5	Always	Make, Go, Feel

5. Conclusion

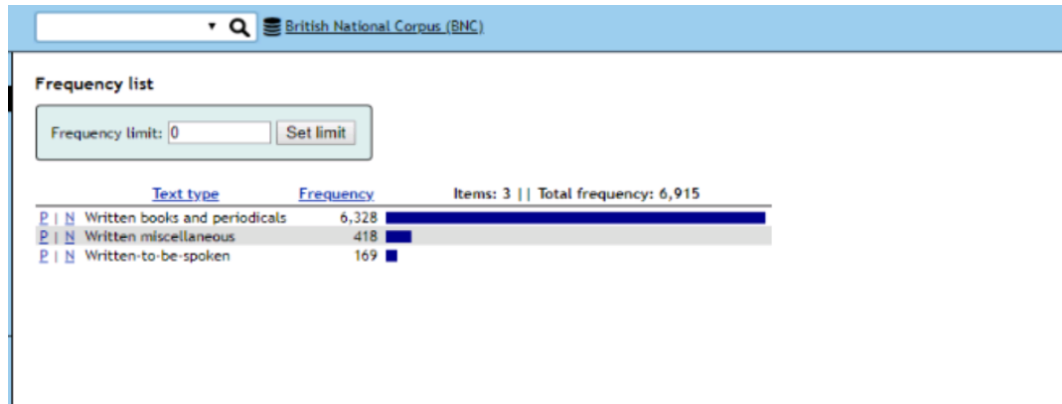
It can be claimed that newspaper corpus (Pak.English-outer circle) has different split infinitives with different frequencies and most importantly with different contexts as compared to BNC (newspaper section-inner circle). The

concept of split infinitives was considered to be limited to spoken genre but this study shows that now it has become an inevitable construction in written genre as well. Along with other studies on the split infinitive, this study shows that the use of split infinitives is widespread and generally accepted across all genres and various levels of formality. To refer to it as “incorrect” or “ungrammatical”, then, seems irrational. The reason behind this increase of split infinitives, especially in academic writing, may simply be a question of semantics. By placing the adverb directly in front of the verb it is supposed to influence, there should be zero chance of ambiguity, which is of utmost importance. In academic writing, there is no room for inadvertent ambiguity. By using Sketch engine, we formed that the use of split infinitives in inner circle and outer circle Englishes both exhibit similarities and differences. In Pakistani English newspaper corpus, ‘immediately’ appears to be the most frequently used breaker in split infinitives under ‘RB’ tag whereas in BNC (Newspaper), the most common breaker is ‘really’. The seemingly contradicting data indicates that the split between inner and outer circle Englishes is not same as initially hypothesized by Kachru (1985). Moreover, by looking at the results, it can be predicted that in future further investigation can be done on split infinitives in terms of different categories e.g. TIME: immediately, temporarily, slowly, simultaneously, DEGREE fully, successfully, properly, effectively, specifically; SCOPE: personally, officially, publicly.

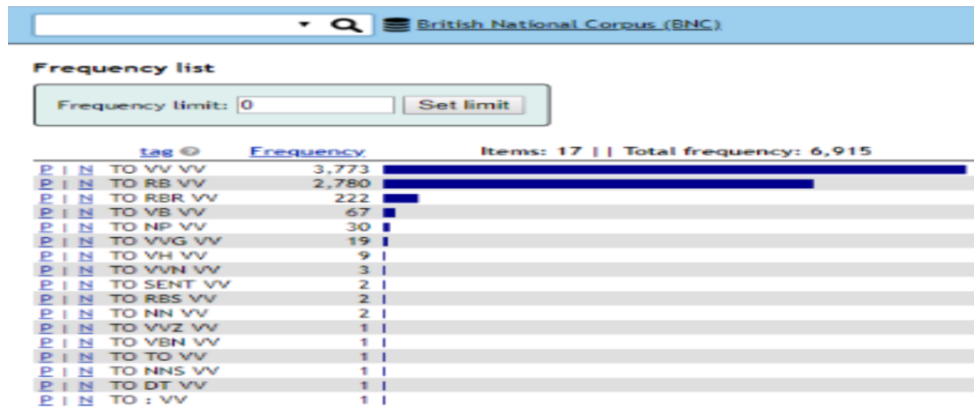
References

- Bao, Z. M. (2005). The aspectual system of Singapore English and the systemic substratist explanation. *Journal of Linguistics*, 41, 237–267.
- Bao, Z. M. (2009). One in Singapore English. *Studies in Language*, 33, 338–65.
- Bautista, M. L. S. (2004). An overview of the Philippine component of the International Corpus of English. *Asian Englishes*, 7, 8–26.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. *TESOL Quarterly*, 34(4), 787-788.
- Evans, S. (2015). Testing the dynamic model: The evolution of the Hong Kong English lexicon (1858- 2012). *Journal of English Linguistics*, 43, 175–200.
- Fitzmaurice, S. M. (2002). Ambiguity in eighteenth century English. A corpus linguistic study of patterns of negation. *Using Corpora to explore linguistic variation*, 9(2), 227-247.
- Martín, C. J., & Barranco, R. J. (2014). On the use of split infinitives in the Asian varieties of English. *Nordic Journal of English Studies*, 13, 130–147.
- Martín, C. J., & García, M. A. (2009). On the use of split infinitives in English. *Language and Computers*, 69(1), 347-364.
- Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics*. Edinburgh, the United Kingdom: Pearson.
- Schneider, E. W. (2003). The dynamics of new Englishes: From identity construction to dialect birth. *Language*, 79(2), 233-281.
- Schourup, L. (1999). Discourse markers. *Lingua*, 107(3-4), 227-265.
- Trudgill, P. (2004). *New dialect formation: The inevitability of colonial Englishes*. Oxford University Press, USA.
- Yule, G. (2010). *The study of language* (4th ed.). Cambridge University Press, United Kingdom.

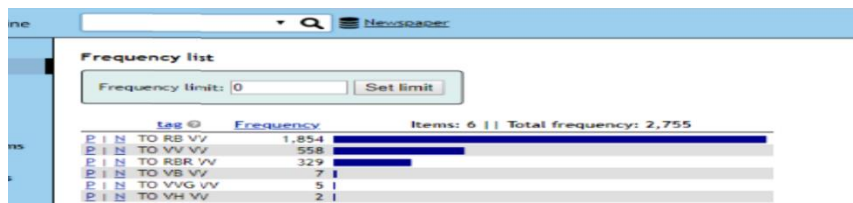
Appendix



1. Text type and frequency of BNC newspaper corpus



2. Frequency of CQL formulas for BNC newspaper corpus



3. Text type and frequency of Pakistani. English newspaper corpus