



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Doctoral Dissertation

*Digital Comments on News:
A Contrastive Study
of the Rhetorical Devices used in
the United Kingdom, India and the United States of America*

Author: Agnija Banerjee

Supervisor: Dr. María Luisa Carrió Pastor

May, 2023

To my beloved husband & family ...

Abstract

In the course of writing, metadiscourse allows a writer to guide, direct and interact with their readers as it inherits the writer’s intention, stance, thoughts, and feelings. The metadiscourse markers allow the readers to comprehend the writer’s judgement and evaluation of the writer’s target audience alongside the idea of the context in use. Hence, metadiscourse has emerged as one of the most popular methods for examining texts and a prevalent research topic since it was first brought to prominence in the early 1980s, continuing to its current popularity.

With the rapid digital progression, nowadays, people are highly interested in digital newspapers, which are the online version of printed newspapers. The prime reason for the attraction towards these digital newspapers is that people get a scope to share their opinions in the form of comments after reading particular news. For this increasing interest of people in digital newspapers, emerging newspapers, as well as well-established and prestigious newspapers, are creating an online version of their printed newspapers. Our study focuses on digital comments (specifically, news comments), where the commenters express their views, opinions, and thoughts through their comments and by replying to the previous comments in the same thread. This doctoral dissertation aims at the exploration and comparison of metadiscourse markers on digital comments of the commenters. It upholds our endeavor of conducting a novel work in the field of language research. Moreover, it can pave the way for future language researchers as well.

In the absence of a benchmark corpus for English news comments, we compiled our corpus by taking into account three popular domains (namely— sports, politics, and entertainment) and two different political ideologies (namely— left-wing and right-wing). We collected 2034 digital comments and a total of 2004 replies from 64 news articles from 12 leading English newspapers of three different countries (the United Kingdom, India, and the United States of America) belonging to three distinct continents (Europe, Asia, and America). In this work, we examine the use of metadiscourse

markers based on four aspects: different domains, English language proficiency, different geo-location, and different political ideologies. Based on these four aspects, this study examines the similarities and contrasts among writers (in our case, commenters), which aids in understanding the writers' usage of metadiscourse markers. In this dissertation, in the context of English digital news comments, for the first time, we present an in-depth quantitative as well as qualitative analysis of metadiscourse markers (interactional and interactive).

We conclude that the commenters use metadiscourse markers differently across different domains; however, the use of metadiscourse may be affected by the commenters' location, language, and political ideologies.

Resumen

Durante la escritura, el metadiscurso permite que un escritor guíe, dirija e interactúe con sus lectores, ya que hereda la intención, la postura, los pensamientos y los sentimientos del escritor. Los marcadores del metadiscurso permiten a los lectores comprender el juicio del escritor y la evaluación de la audiencia objetivo del escritor junto con la idea del contexto en uso. Por lo tanto, el metadiscurso se ha convertido en uno de los métodos más populares para examinar textos y un tema de investigación predominante desde que saltó a la fama a principios de la década de 1980, y continúa con su popularidad actual.

Con la rápida progresión digital, hoy en día, la gente está muy interesada en los periódicos digitales, que son la versión en línea de los periódicos impresos. La razón principal de la atracción hacia estos periódicos digitales es que las personas tienen la oportunidad de compartir sus opiniones en forma de comentarios después de leer una noticia en particular. Debido a este creciente interés de las personas en los periódicos digitales, los periódicos emergentes, así como los periódicos bien establecidos y prestigiosos, están creando una versión en línea de sus periódicos impresos. Nuestro estudio se centra en los comentarios digitales (específicamente, comentarios de noticias), donde los comentaristas expresan sus puntos de vista, opiniones y pensamientos a través de sus comentarios y respondiendo a los comentarios anteriores en el mismo hilo. Esta tesis doctoral tiene como objetivo la exploración y comparación de marcadores metadiscursivos sobre los comentarios digitales de los comentaristas. Sostiene nuestro empeño de realizar un trabajo novedoso en el campo de la investigación del lenguaje. Además, también puede allanar el camino para futuros investigadores de idiomas.

En ausencia de un corpus de referencia para los comentarios de noticias en inglés, compilamos nuestro corpus teniendo en cuenta tres dominios populares (deportes, política y entretenimiento) y dos ideologías políticas diferentes (izquierda y derecha). Recolectamos 2034 comentarios digitales y un total de 2004 respuestas de 64 artículos de noticias de 12 periódicos ingleses líderes de tres países diferentes (Reino Unido, India

y los Estados Unidos de América) que pertenecen tres continentes distintos (Europa, Asia y Estados Unidos de América).

En este trabajo, examinamos el uso de los marcadores del metadiscurso basado en cuatro aspectos: diferentes dominios, el conocimiento del inglés, diferentes geolocalizaciones y diferentes ideologías políticas. Con base en estos cuatro aspectos, este estudio examina las similitudes y los contrastes entre los escritores (en nuestro caso, los comentaristas), lo que ayuda a comprender el uso que hacen los escritores de los marcadores del metadiscurso. En esta disertación, en el contexto de los comentarios de noticias digitales en inglés, por primera vez, presentamos un análisis cuantitativo y cualitativo en profundidad de los marcadores metadiscursivos (interactivos e interactivos).

Concluimos que los comentaristas usan marcadores de metadiscurso de manera diferente en diferentes dominios; sin embargo, el uso del metadiscurso puede verse afectado por la ubicación, el idioma y las ideologías políticas de los comentaristas.

Resumen

Durant l'escriptura, el metadiscurs permet que un escriptor guiï, dirigeixi i interactuï amb els seus lectors, ja que hereta la intenció, la postura, els pensaments i els sentiments de l'escriptor. Els marcadors del metadiscurs permeten als lectors comprendre el judici de l'escriptor i l'avaluació de laudiència objectiu de l'escriptor juntament amb la idea del context en ús. Per tant, el metadiscurs ha esdevingut un dels mètodes més populars per examinar textos i un tema de recerca predominant des que va saltar a la fama a principis de la dècada de 1980, i continua amb la seva popularitat actual.

Amb la ràpida progressió digital, avui dia, la gent està molt interessada en els diaris digitals, que són la versió en línia dels diaris impresos. La raó principal de l'atracció cap a aquests diaris digitals és que les persones tenen l'oportunitat de compartir les seves opinions en forma de comentaris després de llegir una notícia en particular. A causa d'aquest creixent interès de les persones als diaris digitals, els diaris emergents, així com els diaris ben establerts i prestigiosos, estan creant una versió en línia dels seus diaris impresos. El nostre estudi se centra en els comentaris digitals (específicament, comentaris de notícies), on els comentaristes expressen els seus punts de vista, opinions i pensaments a través dels seus comentaris i responen als comentaris anteriors al mateix fil. Aquesta tesi doctoral té com a objectiu l'exploració i la comparació de marcadors metadiscursius sobre els comentaris digitals dels comentaristes. Sosté el nostre afany de realitzar un treball nou en el camp de la investigació del llenguatge. A més, també podeu aplanar el camí per a futurs investigadors d'idiomes.

En absència d'un corpus de referència per als comentaris de notícies en anglès, compilem el nostre corpus tenint en compte tres dominis populars (és a dir, esports, política i entreteniment) i dues ideologies polítiques diferents (és a dir, esquerra i dreta). Recollim 2034 comentaris digitals i un total de 2004 respostes de 64 articles de notícies de 12 diaris anglesos líders de tres països diferents (Regne Unit, Índia i els Estats Units d'Amèrica) que pertanyen a tres continents diferents (Europa, Àsia i Estats Units d'Amèrica).

En aquest treball, examinem l'ús dels marcadors del metadiscurs basat en quatre aspectes: diferents dominis, el coneiximent de la llengua anglesa, geolocalitzacions diferents i ideologies polítiques diferents. En base a aquests quatre aspectes, aquest estudi examina les similituds i els contrastos entre els escriptors (en el nostre cas, els comentaristes), cosa que ajuda a comprendre l'ús que fan els escriptors dels marcadors del metadiscurs. En aquesta dissertació, en el context dels comentaris de notícies digitals en anglès, per primera vegada, presentem una anàlisi quantitativa i qualitativa en profunditat dels marcadors metadiscursius (interactius i interactius).

Concloem que els comentaristes usen marcadors de metadiscurs de manera diferent en diferents dominis; no obstant això, l'ús del metadiscurs es pot veure afectat per la ubicació, l'idioma i les ideologies polítiques dels comentaristes.

Acknowledgements

“Gratitude is the fairest blossom which springs from the soul.”

- Henry Ward Beecher

It is indeed true, as I am genuinely grateful to the Almighty God for everything. I am fortunate enough to have some amazing people in my life, without whom the completion of this doctoral thesis would not have been possible, and for this, I am extremely thankful to them. First and foremost, I want to thank my supervisor Prof. María Luisa Carrió Pastor (Full Professor at the Universitat Politècnica de València), for giving me the opportunity to work with her on such a great topic and for constantly motivating me and providing me the fruitful advice that made this work possible. I am genuinely grateful to Prof. Pastor for introducing me to this topic and helping me in the process whenever I needed her advice and guidance. I want to mention here that not only professionally but also emotionally, when I needed her, she was always there for me, eventually providing me strength to complete the path of my journey. Words are not enough to express my sincere gratitude and immense respect for her.

Secondly, I would like to convey my gratitude to Prof. Giovanni Garofalo (Full Professor at the University of Bergamo, Italy), who provided me with the opportunity of a one-month internship at the University of Bergamo under his esteemed supervision. During the internship, he gave me invaluable advice and support, for which I am truly grateful. I wholeheartedly appreciate his support and help.

Also, I would like to express my gratitude to METOOL, which was developed at the University of Wolverhampton under the research project named “Identification and analysis of metadiscourse strategies in Spanish and in English (IAMET)”.

Next, I would like to express my sincere appreciation to Elena Taulet Grech (International Relations Officer at Universitat Politècnica de València) for encouraging me to start my adventurous journey of Ph.D. Her encouragement gave me the strength to have faith in myself and to plunge into the journey.

I would like to mention that I was fortunate to meet some eminent professors and

researchers during this journey who inspired me a lot. In this regard, I would like to thank Prof. Luisa Chierichetti (Associate Professor at the University of Bergamo), with whom I discussed my research topic.

After that, I would like to thank my previous teachers, professors, and friends from my school, college, and university who motivated me to pursue higher studies. Without a strong foundation, a building cannot exist; likewise, without their help and support, I could not have thought of pursuing a Ph.D. Also, while living in France and Italy, I met some amazing friends who supported and encouraged me emotionally in my journey. In this regard, I would like to thank Anaar, Ana, Alessandra, Elena, and Gaia.

Last but not least, I would like to thank my late grandparents (Pranav Mishra, Amarnath & Neera Banerjee), who are very close to my heart. Also, I would like to thank my family, especially my beloved grandmother (Shukla), my parents (Amit & Mousumi), my twin sister (Megha), my brother-in-law (Supratim), and my relatives (especially my uncles, aunts, cousins, nephews, and nieces) whose constant support, immense love, and encouragement helped me to pursue my Ph.D. I want to give a special thanks to my beloved husband Somnath, who always stood as a pillar for me throughout the entire journey. His unconditional support and encouragement enabled me to complete this doctoral thesis.

Table of Contents

List of Figures	xvii
------------------------	-------------

List of Tables	xix
-----------------------	------------

1 Introduction	1
1.1 Topic and Justification	1
1.2 Hypotheses	5
1.3 Objectives	6
1.4 Research Questions	7
1.5 Structure of the thesis	7
2 English Language in the UK, India, and the USA	11
2.1 Introduction	11
2.2 The origin of the English language in Britain	12
2.3 The origin of the English language in the United States of America	15
2.4 Difference between British English and American English	18
2.4.1 Grammatical differences	19
2.4.2 Difference in pronunciation	20
2.4.3 Difference in vocabulary	22
2.5 World Englishes	24
2.6 The origin of English in India	34
2.7 Characteristics of Indian English	47

2.7.1	Phonetics and Phonology	48
2.7.2	Morphology and vocabulary	49
2.7.3	Syntax	51
2.8	Conclusion	53
3	Metadiscourse	55
3.1	Definition	55
3.2	Classification	57
3.3	Recent Studies	63
3.4	Conclusion	86
4	Digital discourse	87
4.1	Definition	87
4.2	Studies on Digital Discourse	88
4.2.1	Social Media	95
4.2.2	Instant Messaging	98
4.2.3	Comments on News	102
4.3	Conclusion	107
5	Corpus	109
5.1	Introduction	109
5.2	Data Compilation	110
5.2.1	Newspaper selection	112
5.2.2	Newspaper articles selection	113
5.2.3	Comments selection	114
5.2.4	Data pre-processing	116
5.3	Corpus Description	117
5.4	Conclusion	121

6	Method	123
6.1	Introduction	123
6.2	Data Annotation Tool	123
6.3	Data Annotation Process	127
6.3.1	Metadiscourse Markers with Different Senses	129
6.4	Normalization	130
6.5	Conclusion	131
7	Results & Discussions	133
7.1	Introduction	133
7.2	Quantitative Analysis of India Comments	134
7.3	Quantitative Analysis of UK Comments	138
7.4	Quantitative Analysis of USA Comments	143
7.5	In-depth Analysis of Markers	147
7.5.1	Interactional: Hedges	148
7.5.2	Interactional: Boosters	165
7.5.3	Interactional: Attitude Markers	178
7.5.4	Interactional: Engagement markers	192
7.5.5	Interactional: Self-mentions	203
7.5.6	Interactive: Transition markers	210
7.5.7	Interactive: Frame markers	219
7.5.8	Interactive: Endophoric markers	229
7.5.9	Interactive: Code Glosses	237
7.5.10	Interactive: Evidentials	244
7.6	Domain wise Analysis of Comments	251
7.7	Political Ideology Analysis of Comments	256
7.8	Cross-cultural Analysis of Comments	261
7.9	Conclusion	268

8	Conclusions	269
8.1	Answering to research questions	270
8.2	Contributions of the Thesis	281
8.3	Limitations and Recommendation for Further Study	283
	References	287
	Appendix A	307
A.1	Quantitative Analysis of India Comments	307
A.2	Quantitative Analysis of UK Comments	308
A.3	Quantitative Analysis of USA Comments	308
A.4	Domain wise analysis of comments(normalized)	309
A.5	Political Ideology Analysis of Comments (normalized)	310
A.6	Cross-culture Analysis of Comments (normalized)	311

List of Figures

2.1	Three concentric circles of World Englishes proposed by Kachru (1990)	32
6.1	METOOOL: Marker annotation	126
6.2	METOOOL: Marker annotation	127
6.3	METOOOL: Concordance	128
6.4	Analysis of Corpus: snapshots from METOOOL	129
7.1	India: Distribution of <i>Interactive markers</i> (normalized)	134
7.2	India: Distribution of <i>Interactional markers</i> (normalized)	136
7.3	UK: Distribution of <i>Interactive markers</i> (normalized)	140
7.4	UK: Distribution of <i>Interactional markers</i> (normalized)	142
7.5	USA: Distribution of <i>Interactive markers</i> (normalized)	144
7.6	USA: Distribution of <i>Interactional markers</i> (normalized)	146
7.7	Comparisons of the occurrences (normalized) of <i>hedges</i>	164
7.8	Comparisons of the occurrences (normalized) of <i>boosters</i>	177
7.9	Comparisons of the occurrences (normalized) of <i>attitude markers</i>	191
7.10	Comparisons of the occurrences (normalized) of <i>engagement markers</i>	201
7.11	Comparisons of the occurrences (normalized) of <i>self-mentions</i>	208
7.12	Comparisons of the occurrences (normalized) of <i>transition markers</i>	218
7.13	Comparisons of the occurrences (normalized) of <i>frame markers</i>	228
7.14	Comparisons of the occurrences (normalized) of <i>endophoric-markers</i>	234
7.15	Comparisons of the occurrences (normalized) of <i>code-glosses</i>	243

7.16	Comparisons of the occurrences (normalized) of <i>evidentials</i>	250
7.17	domain wise: distribution of <i>interactive markers</i> (normalized)	252
7.18	domain wise: distribution of <i>interactional markers</i> (normalized)	254
7.19	domain wise distribution (normalized) of <i>interactive markers</i> with respect to country	255
7.20	domain wise distribution (normalized) of <i>interactional markers</i> with respect to country	256
7.21	Political ideology based distribution of <i>interactive markers</i> (normalized)	258
7.22	Political ideology based distribution of <i>interactional markers</i> (normalized)	260
7.23	Cross-culture: domain wise frequency of <i>interactive markers</i> (normalized)	263
7.24	Cross-culture: domain wise frequency of <i>interactional markers</i> (normalized)	264
7.25	Cross-culture: political ideology based distribution of <i>interactive markers</i> (normalized)	265
7.26	Cross-culture: political ideology based distribution of <i>interactional markers</i> (normalized)	266

List of Tables

3.1	A model of metadiscourse in academic texts by Hyland (2004, p. 139) .	61
3.2	Study of Hyland (2010, p. 132)	66
3.3	Interactional metadiscourse study by Yazdani et al. (2014, p. 432) . . .	76
3.4	Findings of Kuhi and Mojood (2014, p. 1050)	77
3.5	Findings of Mardani (2017, p. 76)	78
3.6	Wang et al. (2016, p. 79): interactive metadiscourse in the ten reports .	84
3.7	Wang et al. (2016, p. 80): interactional metadiscourse in the ten reports	84
5.1	Overall statistics of the left-wing and right-wing corpus	113
5.2	Left-wing news corpus collection	115
5.3	Right-wing news corpus collection	116
5.4	Left-wing: corpus statistics	118
5.5	Domain-wise statistics of the left-wing corpus	119
5.6	Right-wing: corpus statistics	119
5.7	Domain wise statistics of the right-wing corpus	120
5.8	Corpus statistics based on tokens	121
7.1	Normalized frequencies of <i>hedges</i> found in India, the UK and the USA .	149
7.2	Top-5 (Normalised frequency per 10,000 words) words of lexico-grammatical categories of <i>hedges</i>	154
7.3	Normalized frequencies of <i>boosters</i> found in India, the UK, and the USA	166

7.4	Top-5 Normalised frequency per 10,000 words of lexico-grammatical categories of <i>boosters</i>	170
7.5	Normalized frequencies of <i>attitude markers</i> found in India, the UK, and the USA	179
7.6	Top-5 (normalized frequency per 10,000 words) words of lexico-grammatical categories of <i>attitude markers</i>	183
7.7	Normalized frequencies of <i>engagement markers</i> found in India, the UK, and the USA	193
7.8	Top-5 (normalized frequency per 10,000 words) words of lexico-grammatical categories of <i>engagement markers</i>	196
7.9	Normalized frequencies of <i>self-mentions</i> found in India, the UK and the USA	203
7.10	Top-5 (normalised frequency per 10,000 words) words of lexico-grammatical categories of <i>self-mentions</i>	205
7.11	Normalized frequencies of <i>transition markers</i> found in India, the UK, and the USA	211
7.12	Top-5 (normalized frequency per 10,000 words) words of lexico-grammatical categories of <i>transition markers</i>	213
7.13	Normalized frequencies of <i>frame markers</i> found in India, the UK, and the USA	222
7.14	Top-5 (normalised frequency per 10,000 words) words of lexico-grammatical categories of <i>frame markers</i>	228
7.15	Normalized frequencies of <i>endophoric markers</i> found in India, the UK, and the USA	232
7.16	Top-5 (normalized frequency per 10,000 words) words of lexico-grammatical categories of <i>endophoric markers</i>	235
7.17	Normalized frequencies of <i>code-glosses</i> found in India, the UK, and the USA	239

7.18	Top-5 (normalised frequency per 10,000 words) words of lexico-grammatical categories of <i>code glosses</i>	242
7.19	Normalized frequencies of <i>evidentials</i> found in India, the UK, and the USA	246
7.20	Top-5 (normalised frequency per 10,000 words) words of lexico-grammatical categories of <i>evidentials</i>	249
A.1	India: frequency of <i>interactive</i> markers (normalized results)	307
A.2	India: frequency of <i>interactional</i> markers (normalized)	307
A.3	UK: frequency of <i>interactive</i> markers (normalized)	308
A.4	UK: frequency of <i>interactional</i> markers (normalized)	308
A.5	USA: frequency of <i>interactive</i> markers (normalized)	308
A.6	USA: frequency of <i>interactional</i> markers (normalized)	308
A.7	Domain wise frequency of <i>interactive</i> markers with respect to political ideology	309
A.8	Domain wise frequency of <i>interactional</i> markers with respect to political ideology	309
A.9	Domain wise frequency of <i>interactive</i> markers with respect to country	309
A.10	Domain wise frequency of <i>interactional</i> markers with respect to country	309
A.11	Based on Political ideology: frequency of <i>interactive</i> markers (normalized)	310
A.12	Based on Political ideology: frequency of <i>interactional</i> markers (normalized)	310
A.13	Cross-cultural: domain wise frequency of <i>interactive markers</i> (normalized)	311
A.14	Cross-cultural: domain wise frequency of <i>interactional markers</i> (normalized)	311
A.15	Cross-cultural: political ideology wise frequency of <i>interactive markers</i> (normalized)	311
A.16	Cross-cultural: political ideology wise frequency of <i>interactional markers</i> (normalized)	312

Chapter 1

Introduction

1.1 Topic and Justification

“The newspaper is a greater treasure to the people than uncounted millions of gold.”

- Henry Ward Beecher

The quote mentioned above by Henry Ward Beecher is highly worthy as, for many centuries, newspapers played an important role in serving information to the people about the important events that were taking place around the world. Similarly, with the digital progression and advancement of technology, the internet has become an integral part of the lives of human beings, and it plays a crucial role in bringing the macrocosmic world into the hands of people. Nowadays, people use the internet in every sphere of their lives. Although newspapers have played an essential role in serving information to the people about important events for many centuries, with the digital progression, the internet plays a vital role in serving information to the people. In the present scenario, people are highly interested in digital newspapers, which are the online version of printed newspapers. Surprisingly, for the increasing interest of people in digital newspapers, not only emerging newspapers but also well-established and prestigious newspapers are creating an online version of their printed newspapers. The main reason for attraction for these digital newspapers is that people get a scope to

share their opinions in the form of comments after reading particular news. According to Gerhards and Schäfer (2010), news websites and online mass media play a crucial role in grabbing the attention of a larger audience on various relevant issues. The interaction of the readers not only confines themselves to the comments but also to replying to the posted comments and showing support to those comments using thumps up or down.

The interactive process of the readers is responsible for the emerging craze in digital newspapers where the authors do not reply, but the readers interact by sharing their reactions. The reply of the readers on certain news is important as their views can be biased, unbiased, or neutral, followed by positive and negative comments. Apart from that, digital newspapers are easily accessible with the help of the internet on various electronic gadgets such as computers, mobile, tab, etc., which helps people to get information anywhere around the globe. Digital newspapers provide news to the readers either free of cost, or they have to pay a subscription fee.

According to Hyland and Jiang (2022), metadiscourse has emerged as one of the most popular methods for examining texts, as evidenced by 29,500 publications containing the term ‘metadiscourse’ that is returned by a Google Scholar search, 11,000 of which were published in the recent five years. From the last decade, researchers have investigated metadiscourse markers in various genres which includes academic writing (Hyland 2010, Alshahrani 2015, GezegİN and Melike 2020), persuasive writing (Noorian and Biri 2017, Korau and Aliyu 2020), research articles (Gillaerts and Van de Velde 2010, Sultan 2011, Cao and Hu 2014, Salek 2014), writing of the EFL learners (Jalilifar and Alipour 2007, Rad 2020), news reports (Boshrahadi et al. 2014), news articles (Yazdani et al. 2014), news editorials (Kuhi and Mojood 2014), opinion columns (Dafouz-Milne 2008), opinion writings on online platforms (Biri 2018), travel vlogs (Huang et al. 2020), advertising (Fuertes-Olivera et al. 2001) etc.

As this doctoral dissertation deals with digital comments (specifically, news comments), it is crucial to understand how language is perceived and used by speakers

and writers. Pragmatics refers to the way people use language. It also deals with the context in which a language is used. According to Archer et al. (2013), in 1938, the eminent philosopher Charles Morris first used the word ‘pragmatics’ to describe the theory of signs where it was found in the tripartite division of semiotics that includes syntax, semantics, and pragmatics. The authors defined the term pragmatics as “the relationship between signs and their interpreters”(Archer et al. 2013: p. 3). In Linguistics, pragmatics is referred to various aspects such as phonetics, syntax, semantics, lexicology, etc. According to Crystal (1987, p. 120), “pragmatics studies the factors that govern our choice of language in social interaction and the effects of our choice on others.” Later, Crystal (2011) stated that, in linguistics study, pragmatics is concerned with the use of language from the users’ point of view. The study of pragmatics focuses on the problems that the users face while interacting in various social situations and the context of the act of communication. It is also concerned with the effects of the user’s language on the other participants. Scholars extensively study it as it is the medium of conveying meaning in verbal as well as non-verbal communication. Yazdani et al. (2014) stated that meaning is not only restricted to the simple interpretation of the semantic and syntactic basis of a sentence, but also there is something that is beyond meaning. Similarly, Finch (2000) expressed his view by stating that Pragmatists concentrate on what is implicit in statements and how we perceive them in different settings. They are more concerned with the force of what is communicated by an utterance than with the meaning of what is uttered, which is to say, with the manner and style of a speech. In the past, a number of researchers, such as Sperber and Wilson (2002), Wharton (2003), and Wilson (2005), have studied pragmatics extensively.

Our study focuses on digital comments (specifically, news comments), where the commenters express their views, opinions, and thoughts through comments and replying to the previous comments in the same thread. This doctoral dissertation provides a short description of pragmatics to understand the relationship between the reader and the writer (in our case, commenters). Here, we would like to mention that our pri-

mary focus is on metadiscourse, specifically interactional metadiscourse that engages the readers in the text and interactive metadiscourse that guide the readers throughout the text.

However, to the best of our knowledge, while writing this dissertation, no such study has been carried out so far on newspaper comments written in English. Moreover, we found that any study has not been carried out so far on the use of metadiscourse that encompasses the use of metadiscourse markers in the context of different cultural backgrounds, such as different countries, different speakers, etc. So, there is a scope to extend the use of metadiscourse devices in newspapers regarding comments by commenters of different nations, continents, or speakers of English as native or non-native. This motivates us to pursue this research work, where we intend to show a contrastive analysis of the metadiscourse markers of the news comments from three countries (the United Kingdom (from now on UK), India, and the United States of America (from now on USA)) belonging to three different continents (Europe, Asia, and America).

This doctoral dissertation puts an emphasis on the comments along with the replies of the readers in digital newspapers. It covers three contexts, namely, European, Asian, and American, by focusing on the comments of the readers based on three domains, namely, politics, sports, and entertainment. For this study, we selected two categories of newspapers, namely, left-wing and right-wing newspapers. The commenters on news were chosen from the left-wing and right-wing newspapers of three nations: the UK, India, and the USA. There were national as well as international news that were collected from the three domains: politics, sports, and entertainment.

As for the significance or contributions of this doctoral study, we would like to mention the following aspects:

- This study could be referenced as a pioneering work on English news comments.
- The compiled corpus that was created for this dissertation, in our opinion, will serve as a linguistic resource for future study.

-
- The employment of rhetorical devices, particularly metadiscourse markers, by commenters from the three countries (UK, India, and USA) that belongs to three continents (Europe, Asia, and America) is covered in this dissertation. Therefore, this study demonstrates the similarities and contrasts among a variety of writers (in our case, commenters) from different countries belonging to different continents, which aids in understanding the writers’ cultural backgrounds.
 - This study examines how commenters from three countries—the UK, India, and the USA—supporting two different political ideologies (left-wing and right-wing) employ metadiscourse markers and identifies similarities and contrasts in their use.
 - From this study, a perspective on the use of metadiscourse markers (interactional and interactive) by the commenters from three different domains (entertainment, politics, and sports) could be comprehended.
 - Furthermore, in this study, we investigated English news comments from non-native speakers of the language from India as well as native English speakers from the UK and the USA. Our investigation could be of immense importance in understanding how both native and non-native English speakers use metadiscourse markers.

1.2 Hypotheses

The following hypotheses are taken into account in accordance with the research objectives of this doctoral dissertation:

- **H1:** *The usage of metadiscourse markers can be used to identify the writers’ cultural backgrounds.*

In this research work, we hypothesize (H1) that people who write for digital newspapers—especially those who write comments—and who reside in three different

nations on three distinct continents tend to employ metadiscourse markers in different ways.

- **H2:** *There are differences in the use of metadiscourse markers among the writers who comment on digital newspapers while they comment on the news of different political ideologies or different domains.*

The intuition of hypothesis H2 is that even the commenters from the same cultural background have a tendency to employ metadiscourse markers in different ways when they are commenting on news from different political ideologies (such as left-wing, right-wing, etc.) and domains (e.g., sports, politics, entertainment, etc.).

1.3 Objectives

As a general objective of this doctoral dissertation, we consider analyzing digital comments on specific news written in English by commenters from different nations.

The specific objectives are as follows:

- To identify and analyze metadiscourse markers used by the commenters of three nations (the United Kingdom, India, and the United States of America).
- To identify the similarities and differences in the use of *interactional* and *interactive* metadiscourse markers used in the digital comments by the commenters of the UK, India, and USA.
- To detect the similarities and differences in the use of metadiscourse markers while commenting on digital news of diverse domains by the commenters.
- To present the similarities and differences in the use of metadiscourse markers by the commenters supporting different political ideologies.

1.4 Research Questions

Considering diverse domains, English native and non-native speakers, geo-location, and political ideology, we aim to answer the following research questions in this doctoral dissertation:

- **RQ1** *Are there any differences or similarities in how writers (in our case, commenters) use metadiscourse markers across politics, sports, and entertainment domains?*
- **RQ2** *What are the similarities or differences between English native speakers and non-native speakers who comment on digital newspapers in terms of the use of metadiscourse markers?*
- **RQ3** *Do commenters from the UK, India, and the USA employ metadiscourse markers similarly or differently depending on their geo-location and culture?*
- **RQ4** *What are the similarities or differences in the use of metadiscourse markers when the commenters support a political ideology—left-wing or right-wing?*

1.5 Structure of the thesis

This doctoral dissertation is structured considering the above-mentioned objectives and research questions. It comprises eight chapters in total. We begin with Chapter 1 i.e., the Introduction, that initially discusses the topic in general and mentions the reasons and justification for selecting this topic for our dissertation. In addition, we provide a brief discussion of pragmatics that deals with language and its use in context by the people, which, eventually, in this study, helps to understand the relationship between the reader and the writer (in our case, we would refer to them as commenters). We also mention the factors that motivated us to pursue this doctoral study. Section 1.2 upholds the hypothesis that we consider for this study; the general, as well as specific

objectives, are mentioned in Section 1.3, and the research questions are discussed in Section 1.4. Finally, we conclude the first chapter by describing the structure of the thesis.

In Chapter 2, we discuss the birth of the English language in Europe in Section 2.2. The origin of the English language in the United States of America is mentioned in Section 2.3. In Section 2.4, we describe some of the differences between British and American English, including grammatical differences, the difference in pronunciation, and differences in vocabulary. Section 2.5 describes World Englishes, following Section 2.6, which describes the origin of English in India. We mention some of the characteristics of Indian English in Section 2.7, including phonetics and phonology, morphology and vocabulary, and syntax. In Section 2.8, we conclude the chapter by mentioning different sections.

Chapter 3 presents the theoretical background of metadiscourse. In this dissertation, metadiscourse plays a pivotal role as this study mainly focuses on the similarities and differences in using metadiscourse markers on digital comments. This chapter starts with the definition of metadiscourse in Section 3.1, which follows the classification of metadiscourse in Section 3.2. The recent studies carried out on metadiscourse are mentioned in Section 3.3. Section 3.4 concludes this chapter by stating the research works carried out so far on metadiscourse.

Chapter 4 includes the theoretical background of digital discourse, with the definition in Section 4.1. The literature on digital discourse is discussed in Section 4.2. Subsections 4.2.1 and 4.2.2 briefly explain social media (Twitter, Facebook, and Instagram) and instant messaging (WhatsApp, Telegram, and Viber), as these are also a part of digital discourse. Subsection 4.2.3 describes comments on news and the literature on news comments. Finally, this chapter concludes in Section 4.3.

Chapter 5 mentions the corpus with the introduction in Section 5.1; Section 5.2 describes data collection, including newspaper selection, newspaper articles selection, comment selection, and refining the data. Next, Section 5.3 provides a detailed de-

scription of the corpus. We conclude this chapter in Section 5.4.

Chapter 6 demonstrates the research method to conduct this research. Section 6.2 describes the data annotation tool used in this study to conduct the research. Here, we also mention the steps followed while annotating a corpus with METOOL (the annotation tool). Section 6.3 describes the data annotation process. In subsection 6.3.1, we discuss metadiscourse markers with different senses, along with one example from our corpus. Next, Section 6.4 talks about the normalization process following the conclusion of the chapter in Section 6.5.

Chapter 7 presents the results of this study. With the introduction in Section 7.1, the quantitative analysis of the Indian comments is portrayed in Section 7.2, following the quantitative analysis of the UK and the USA comments in Sections 7.3 and Section 7.4, respectively. Section 7.5 presents the in-depth analysis of the interactional and interactive metadiscourse markers. The analysis was carried out taking into account the quantitative and qualitative use of the markers by the commenters. Section 7.6 depicts the domain wise analysis of the comments. Section 7.7 portrays the analysis of the comments based on the political ideology of the commenters. The cross-cultural analysis of the comments is shown in Section 7.8. Finally, in Section 7.9, the conclusion of this chapter is drawn.

Chapter 8 provides the conclusion of this doctoral dissertation. In Section 8.1, we summarize this study's research findings and observations by answering the research questions that were formed at the beginning of this chapter (Section 1.4). In Section 8.2, the contributions of this doctoral study are portrayed. Section 8.3 outlines this study's limitations and mentions some future directions that can be further explored.

Chapter 2

English Language in the UK, India, and the USA

2.1 Introduction

English is one of the most spoken languages in the world, with more than a billion speakers from various nations. Nowadays, English is spoken almost in every continent and is pursued as a second language by an enormous number of speakers from around the world. [Algeo et al. \(1992, p. 1\)](#) stated that “the history of a language is intimately related to the history of the community of its speakers, so neither can be studied without considering the other”. In order to understand how English became the most popular and common language in the world, it is crucial to understand its historical development that includes its speakers. The other factors include geographical location, wars, political situation, culture and traditions, government and administration, customs and rituals, etc. The English that we speak today has come through a great historical development throughout the ages, and thus the whole development process can be categorized into three periods: Old English, Middle English, and Modern English. It is mentioned in [Blake \(1996\)](#) that the division of the history of the English language into three major periods: Old, Middle, and Modern was first suggested by

Henry Sweet while delivering a lecture in 1873. The written version of Henry Sweet includes “I propose, therefore, to start with the three main divisions of Old, Middle, and Modern, based mainly on the inflectional characteristics of each stage. Old English is the period of full inflections (*nama, gifan, caru*), Middle English of leveled inflections (*namme, given, caare*), and Modern English of lost inflections (*naam, gic, caar*)” (Blake, 1996, p. 6). He mainly categorized the periods on the basis of sound, and it is highly acceptable by other scholars as well.

2.2 The origin of the English language in Britain

Though people argue on the fact when English began and how it came into existence, looking at the historical evidence, it can be assumed that the English that we speak today is different from what it was used in the early stages of the language, and also it has gone through a great shift and change to attain its present form. According to Kirkpatrick (2010), Old English can be regarded as the offspring of Proto-Germanic, which itself comes from the Proto-Indo-European group of languages. Old English has some common features with Gothic, Old Norse, and Old High German. It is stated in Kirkpatrick (2010) that Proto-Germanic refers to the hypothetical parent language that is re-established in accordance with the early surviving texts in the Germanic daughter languages. Again, Proto-Indo-European refers to the hypothetical parent language that is developed in accordance with the early surviving texts in all of the Indo-European languages. Later, various sound shifts and changes in the Proto-Germanic language resulted in the prominence of the daughter Germanic languages, among which Old English is noteworthy to mention. Archaeological evidence shows that the speakers of Germanic languages were from Denmark and Sweden, who later came in contact with the Romans due to trade and mercenary purposes. The spread of the Germanic people, along with the connection with Romans, indicates the Roman influence in almost all the sectors of the life of the Germanic people, including trade and business, commerce,

agriculture, legal institutions, etc. Due to this, the process of borrowing loan words started, which indicates the adoption of a particular word from another language with a slight variation or modification. The borrowing of loan words continued during the invasion of the Romans.

Before the decline of the Roman Empire, the inhabitants of Britain were the Romanized Celts, who were the speakers of the Brittonic branch of Celtic (Kirkpatrick, 2010). In the fifth century, the Romans left Britain and paved the way for the invaders. During this time, the Irish, Scots, and Picts from Ireland and Scotland came to Britain and began to occupy the lands of the Romanized Celtic people. After that, the Anglo-Saxons arrived in Britain as mercenaries to help the-then inhabitants of Britain deal with the Picts and Scots (Albert and Stone, 1979). Then gradually, they began to settle there, and nearly at the end of the fifth century, they settled permanently in Britain. According to Blake (1996), in the fifth century AD onwards, the Germanic people unitedly known as the Anglo-Saxons came to Britain and brought the English language with them, which was the West Germanic branch of language. Gradually it spread around the British Isles and also arrived in parts of Wales, Cornwall, Cumbria, and Southern Scotland. Blake (1996) argues that it would be inappropriate to state that the English that we speak today was there at the time of the arrival of the Anglo-Saxons. Before their arrival, the tribal groups of Britain have their own dialects and followed a single monarch. But with the Anglo-Saxons, a variety of English belonging to the West Germanic branch arrived in Britain. In 597, after the arrival of St. Augustine in Britain, the English kings and their subjects were converted to Christianity, and the borrowing of loan words from Latin was continued. During this time, Latin was used to perform all the religious and administrative services. Then the establishment of monasteries and schools was started, and along with that, the copying of the biblical and Latin texts was also started. At this time, the borrowing of Latin words into English was done to a large extent (Blake, 1996). Later, the arrival of the Scandinavian people in Britain in search of wealth took place, which led to the borrowing of words

from the Old Norse.

The beginning of the Middle English period in Britain was in 1066 with the Norman Conquest. During this time, there were three prime languages in Britain: Latin, French, and English. It became a trilingual society where Latin was used for administrative and ecclesiastical purposes, French was the language used by the people of the upper class in Britain, and English was the language for communication of the lower class people of the society. It was mainly the language of the peasants and the common people. During the twelfth and thirteenth centuries, a considerable amount of loan words were borrowed from the French. However, the situation changed in the fourteenth century, when the popularity of French declined, and English was used in education and also by the upper-class people (Kirkpatrick, 2010).

In 1485, the Middle English period ended with the rise of the Tudor dynasty. After this period, various events took place which have great significance in the linguistic development of the English language. One of the most important events was the establishment of the printing press. Instead of copying the texts by hand, it was far easier to get the texts in printed form and also in huge numbers. However, the crucial concern for the authors was the variation in the dialects. In this situation, the authors followed ‘the most common or understandable of several variant forms’ (Kirkpatrick, 2010, p. 27). As London was the main city of Britain, it played a crucial role in the evolution of standard English. Keene (2000) stated that London is probably credited with having a significant impact on the development of Standard English, but less as a center of power and government and more as a hub for exchange and communication.

During this period, in English, most of the loan words were mainly borrowed from Latin. After Latin, most of the loan words were borrowed from French, Italian, Spanish, and Dutch. Also, there were some Arabic, African, Hindi, Urdu, and Tamil words that entered English during this period. Another crucial change that took place starting from the Middle English period and continuing to the Modern English Period was the Great Vowel Shift. During the period of 1400-1700, the Great Vowel Shift took place,

which refers to an enormous change in the pronunciation of English words. This change includes the change in the pronunciation of the vowel sounds of English, where formerly, the vowels were pronounced in one particular place in the mouth. However, later due to the Great Vowel Shift, the vowels were pronounced in the upward direction of the mouth. This change basically affected the pronunciation of the long vowels. According to Blake (1996), during the seventeenth century, several books and pamphlets were published in ‘correct’ English that later created an urge among the compilers to prepare a list of acceptable words used in English. At this time, the concept of dictionaries came forward, and in 1755, Dr. Samuel Johnson’s dictionary was marked as one of the most dominant dictionaries of the English language. A notable change was noticed in English grammar, where the grammatical rules were mainly transformed from Latin. During this period, the main concern was to establish a standard written English language. With the publication of the Lyrical Ballads in 1789, the whole concept of the standardization of the English language was changed. Lyrical Ballads is a collection of poems by William Wordsworth and Samuel Taylor Coleridge that is considered as a milestone in the history of the English language and literature. Instead of using formal and ornamented language in poetry, Lyrical Ballads first manifested the idea of using earthly, mundane, day-to-day chores and activities of the common people. The prime concern of the poets was to make poetry easily accessible to the common people.

2.3 The origin of the English language in the United States of America

In the nineteenth century, the prime focus was on the diversity of the languages, and publications started to be in the non-standard variety of the English language. After the First World War, the United States of America became a global power, and English was used as the native language. However, the English that is used in the United States is different from the English spoken in the United Kingdom. During the

nineteenth century, the movement of English started with voyages to different countries and continents.

Crystal (1997) stated that the first failed attempt of the voyage from England to the 'New World' was in 1584 by Walter Raleigh. Later, a group of people from England settled near Roanoke Island (presently called North Carolina) and clashed with the local inhabitants of that place. The settlement actively started in 1607 when a group of settlers came and settled in the Chesapeake Bay. These explorers named the settlement after James I as Jamestown. They named that place Virginia after the 'Virgin Queen' Elizabeth. Gradually, people began to settle in the coastal areas, and in 1620, a group of Puritans appeared who could not reach Virginia due to extreme weather conditions and settled in Cape Cod Bay (presently known as Plymouth, Massachusetts). Later, this group of settlers was known as the 'Pilgrim Fathers'. This group of settlers belongs to different origins and backgrounds. The common thing that made them stay connected with each other was a dire need to find land to settle permanently where they could follow their own religion and religious practices following the Bible. This settlement was regarded as one of the most successful settlements, and during 1640, a vast number of people came to that place to settle down. During this time, there were two prime settlements: one was in Virginia and the other was in Plymouth. Though the settlers belong to different linguistic backgrounds, the settlers of Virginia were from Somerset and Gloucestershire, and the settlers of Plymouth were from the eastern countries of England mainly from Lincolnshire, Nottinghamshire, Essex, Kent and London, etc. (Crystal, 1997). A notable difference was seen in the pronunciation between these two groups of settlers. The immigrants continued to come, and they moved to the west, south, and mid-western area and settled there. As a result, the variation in the dialects can be noticed among the settlers. Crystal (1997) mentions that in the seventeenth century, a group of immigrants came from mainly the Midlands and northern England, and along with them they bring a new variety of English. These people were known as Quakers who later settled in Pennsylvania. In this way, new linguistic varieties were

introduced by the immigrants. The differences among the regional dialects become to fade gradually when the settlers living alongside each other and speaking a different variety of English settled in New York. An enormous number of immigrants from Ireland and Scotland were seen to come to America in the eighteenth century. These Scots and Irish settlers moved along the coast of Philadelphia and moved to the mountains. In 1776, one of the seven colonial populations was occupied by the Irish and Scots. These Irish-Scots came with an accent that has a vast similarity with present-day American English. Later, Spanish, French, Dutch, and German people also settled in the western and northern parts of the country.

It has been a matter of argument regarding the first settlement of the English people in America and how the English language developed throughout the years. [Kovecses \(2000\)](#) mentioned that the first successful English settlement was in Jamestown in 1607. However, it is argued in [Dillard \(2014\)](#) that with the expedition of John Cabot in 1497, the first officially recognized English-speaking group came to America. In describing the advent of American English, the importance of exploration and colonization cannot be denied. Additionally, the languages spoken by the Native Americans and the immigrants were also important in the development of American English.

In 1619, the slaves from Africa were brought and they were taken first to Jamestown. With the legalization of slavery in 1650, the importation of slaves from Africa continued till the middle of the nineteenth century. This period is marked by some scholars as one of the most linguistically crucial periods as during time, the speakers of what we call now American English emerged on the continent of North America. The settlement of the Spanish was seen mainly in the southern part of the continent. The French and Dutch people were found to be settled in the New York region. [Kovecses \(2000\)](#) stated that the Atlantic seaboard consisted of three areas ranging from the north to south, namely, New England, the Middle Atlantic States, and the South Atlantic States. The New England area, comprised of a group of settlers, established their colony in Plymouth, Massachusetts. Later, the settlement continued in Connecticut,

Rhode Island, and New Hampshire. The Middle Atlantic States comprised of what is presently called New York and the settlers were Dutch. The settlement in the South Atlantic colonies consisted of Virginia which was later followed by North Carolina, South Carolina, and Georgia. The settlement continued after the end of the War of Independence, and settlers were seen to move towards and beyond the Mississippi River. The new settlers who selected the United States as their home were mainly from Ireland and Germany who further discovered the unexplored territories. During this period, English was regarded as the national language of the new country. This period is very crucial for the development of American English. In the time of the nineteenth century, immigration still continued, and millions of immigrants were seen from different parts of the world, mainly from Southern and Eastern Europe. Later, the immigrants from Mexico, Puerto Rico, and Cuba were also seen to be settled in different parts of the United States. During this time, a tendency of using the mother tongue in English started. Also, during this time, an upsurge of unity was seen among the immigrants in accepting English as their common language despite the cultural diversification (Crystal, 1997). Later, the English language gradually spread in Canada, New Zealand, Australia, Asia, and all over the world and it has become the most commonly spoken language in most of the countries as well.

2.4 Difference between British English and American English

Due to the vast immigration, a lot of changes in pronunciation, vocabulary, and grammar were seen in American English. Kovecses (2000) hinted at the point that various qualities of seventeenth and eighteenth century English are still prevalent in American English, whereas they are not in use in British English anymore. The differences include changes in pronunciation, word spelling, number expressions, idioms, the difference in vowel and consonant pronunciation, etc. In this study, we focus on the

differences regarding three aspects: grammar, pronunciation, and vocabulary.

2.4.1 Grammatical differences

Verbs: There are various grammatical differences between British and American English. In British English, the past tense of the word ‘get’ is ‘got’ while the Americans still use the past participle form ‘gotten’. In British English, the past participle form of the verb ‘get’ is not used after the seventeenth century. A good example of this could be found in the study of [Zhang and Jiang \(2008, p. 70\)](#): a) John has got much better during the last week (British English). b) John has gotten much better during the last week (American English). In American English, the verb ‘got’ refers to owning something.

Tense: The difference can be seen in the use of tense. In British English, the past perfect tense is used while describing an incident that just happened recently, while in American English, simple past tense is used in these kinds of situations. For example, a) I’ve just seen your brother (British English). b) I just saw your brother (American English) [Zhang and Jiang \(2008, p. 71\)](#).

Preposition: In British English and American English, the use of prepositions is different in most cases. While constructing the phrases, the difference can be noticed between British and American English. For example, These dresses are in a sale. (British English). These dresses are on sale. (American English). While in British English, prepositions are often used in a sentence, in American English, prepositions are often omitted. For example, I will see you on Monday. (British English). I will see you Monday. (American English) [Zhang and Jiang \(2008, p. 70\)](#).

Collective noun: Another difference is seen in the use of collective nouns in British and American English. While in British English, the collective nouns are used with a plural verb, the singular verb is used in American English. For example, the plural verb ‘are’ is used for the collective noun ‘government’ in British English. On the contrary, the singular verb ‘is’ is used for the collective noun ‘government’ in American English

(Kovecses, 2000).

Articles: There are differences between British English and American English in the use of articles. In British English, 'the' is mostly used in standard expressions, while in American English, 'the' is not used frequently. In British English, 'the' is used before expressions like 'all the afternoon', 'all the week.' On the contrary, in American English, 'the' is not used before these expressions: 'all week', 'all afternoon' Zhang and Jiang (2008, p. 72). Again, there are some expressions in British English where articles are not used, such as 'in hospital', 'at university' etc. In American English, the article 'the' is used, such as 'in the hospital', 'at the university'. There is a difference in the position of the articles 'a' and 'an' with the word 'half'. In British English, the articles (a, an) are used after the word 'half', for example, 'half an hour', 'half a dozen' etc. In American English, the position of the article is before the word 'half', such as 'a half hour', 'a half dozen' etc. Zhang and Jiang (2008, p. 72).

Pronouns: There are differences in the use of pronouns also. The indefinite pronoun 'one' is used differently in British and American English. In British English, 'one' is repeatedly used in a sentence, while in American English, it is often replaced by another pronoun. For example, One cannot succeed unless one tries hard (British English). One cannot succeed unless he tries hard (American English) Zhang and Jiang (2008, p. 72).

Adverbs: In British English, adverbs are placed mostly after the auxiliary verb, whereas in American English, adverbs can be placed before or after the auxiliary verbs without changing the meaning of the sentence. For example, a) They never will agree to it, or b) They will never agree to it (American English). They will never agree to it (British English) Zhang and Jiang (2008, p. 73).

2.4.2 Difference in pronunciation

A vast amount of difference is noticed in the pronunciation of British and American English. There is a difference in the pronunciation of the rhotic accent that refers to the pronunciation of the letter 'r' after a vowel within a syllable. Whereas American

English is rhotic, British English is mostly non-rhotic. The rhotic accent is produced as a retroflex approximant in English. While producing the rhotic accent, the tongue moves towards the gum, and the tip of the tongue is curled back towards the roof of the mouth. The term retroflex approximant defines that during the production of the sound, the tongue is drawn back into the mouth, which describes the term retroflex. Again, it is called approximant because the tip of the tongue does not touch the gum, so no friction is produced, and the vocal tract remains open. The examples of words with rhotic sounds are ‘water’, ‘later’, ‘born’, ‘birth’, ‘heard’, ‘door’, ‘university’ etc. (Zhang and Jiang, 2008). Kovecses (2000) mentions that at the end of the eighteenth century, the use of /r/ was shunned as in the words ‘bar’, ‘colo(u)r’. In order to mark the class distinction, the rhotic accent was removed by the upper-class people of England, which was gradually adopted later by the middle-class people of England. On the contrary, the lower class people of Scotland and Ireland did not adopt this change, and they remained rhotic (Gomez, 2009).

There are differences in the pronunciation of vowels as well as consonants in British and American English. The pronunciation of the sound /u/, as in words like ‘new’ and ‘knew’, is different. In British English, it is pronounced as /ju/. In American English, the pronunciation was kept with less /j/ while it was prominent in British English. In American English, the pronunciation of ‘short o’ as in words ‘hot’, ‘top’, ‘lot’ etc., is flat and not rounded. In British English, this sound is pronounced as rounded. During the eighteenth century, the pronunciation of the flat ‘o’ is discarded in England while it is still prevalent in America (Kovecses, 2000). In British and American English, the stress pattern also works differently in various situations. For example, in American English, in words like ‘necessary’ and ‘secretary,’ the stress is given on the last syllable. On the other hand, in British English, the stress is given on the first syllable (Kovecses, 2000). Also, differences can be noticed in the articulation of sounds in British and American English. There is a tendency among Americans to pronounce unstressed syllables. Another crucial difference between British and American English is noticed

regarding the intonation. When the British people speak, their voice could be very high or very low. However, while Americans speak, their voice usually remains flat, and it does not go very high or low.

2.4.3 Difference in vocabulary

There are differences in the use of vocabulary in British and American English. *Kovecses (2000)* stated that the earlier English usage was adopted by the Americans, and they still follow that, whereas the British people have either discarded it or changed it. In American and British English, a particular word can have different meanings. For example, the word ‘mad’ is used to mean ‘insane’ in British English, whereas it means ‘angry’ in American English. Earlier in the Elizabethan era, Shakespeare used the word ‘mad’ to mean angry. But in modern British English, it is used to mean ‘insane’. Another important example that marks the difference between British and American English is the word ‘fall’. While in British English, ‘autumn’ is used instead of ‘fall’ to indicate a particular season, the word ‘fall’ has been used for the same purpose by the Americans till now. In earlier English usage, the word ‘fall’ is used that was later adopted by the Americans (*Kovecses, 2000*).

In American English, the word ‘sick’ means ‘ill’, which was also considered the same as in the seventeenth century in England. In modern British English, the word ‘sick’ means ‘ready to vomit, to feel nauseated’. The use of the noun ‘druggist’, which means ‘a person who makes medicine’, is used in America till now, whereas in modern British English, instead of ‘druggist’, the word ‘chemist’ is used. Similarly, the word ‘apartment’ was used in 1641 in England. But later, in modern British English, the word ‘flat’ is used instead of ‘apartment’ to mean ‘a number of rooms in a house or building that belong to a single family’. But in American English, the word ‘apartment’ is still used *Kovecses (2000, p. 29)*. In modern British English, the noun ‘bug’ means ‘small insect infesting dirty houses and beds’ (*Oxford Advanced Learner’s Dictionary*) *Kovecses (2000, p. 28)*. Earlier, it was meant to refer to any small insect and not

specifically an insect that infects beds and houses. In American English, the former meaning is accepted and used, while the British have discarded it. Again, the difference can be noticed in the vocabulary between British and American English in the use of the word 'bloody'. The meaning of the word 'bloody' is 'covered with blood'. This meaning is also used in America till date. But in British English, this word is associated with negative feelings that express anger. *Kovecses (2000, p. 29)* stated that the use of this word created a disturbance in England when this word is used in G.B Shaw's 'Pygmalion'. There is also a great impact of the French loan words that were later carried out and used in American English as during the exploration of the new continent, a good number of French people moved and got settled in America. For example, the words like 'bureau', 'depot', 'shanty', 'carry-all', 'voyageur' are French words that later came into usage in American English. Apart from French, the influence of Spanish, Dutch, and German is also seen in enriching the vocabulary of American English. For example, the words like 'barbecue', 'chocolate', 'tortilla', 'tomato' are of Spanish origin that later came to be a part of American English vocabulary. Again, there are some Dutch words that were later added to the vocabulary of American English, such as 'cookie', 'landscape', 'boss', 'waffle', etc. Similarly, some borrowed German words are 'hamburger', 'noodle', 'pretzel', etc. *Kovecses (2000, p. 33)*.

There are also spelling differences between American and British English. There are some words that are of Greek origin that consist of 'ae', while in American English, only 'e' is used instead of 'ae'. For example, the word 'gynaecologist' (a skilled doctor who used to treat women-related problems or diseases) is spelled with 'ae' in British English, while the spelling is 'gynecologist' in American English. There are some words that are named differently in British and American English. For example, the word 'biscuit' is used in British English, whereas to mean the same thing in American English, 'cookies' is used. Similarly, in Britain, the word 'chips' is used, and in America, it is 'fries'. The above-mentioned differences show the distinction between British and American English.

2.5 World Englishes

Due to the colonial development around the world, the movement of English expanded in Africa, Asia, and South Pacific region. Later in the twentieth century, this movement took a remarkable turn when English was regarded as the ‘official’ or ‘semi-official’ status in some of the independent countries. According to [Crystal \(1997, p. 30\)](#), “The language has penetrated deeply into the international domains of political life, business, safety, communication, entertainment, the media, and education.” In this context, the concept of ‘world Englishes’ ([Bolton, 2012a](#)) helps to understand the expansion of the English language around the world. It is mentioned in [Bolton \(2012a, p. 13\)](#) that the term ‘world Englishes’ refers to ‘the localised forms of English found throughout the world, particularly in the Caribbean, parts of Africa, and many societies in Asia’.

During the twentieth century, the categorization of English was mainly based on the difference between ‘native’ and ‘non-native’ speakers of English. These categories are as follows: English as a native language (ENL), English as a second language (ESL), English as a foreign language (EFL), and English as an international language (EIL) ([Bolton, 2012a](#)). However, [Halliday et al. \(1964\)](#) stated that English is no longer the exclusive language of the British or even the British and Americans; rather, it has become a universal tongue that a growing number of people use for at least some of their needs. The author also argued that there is no longer a widespread belief in West Africa, the West Indies, Pakistan, and India that the English spoken in England with the received pronunciation (RP) is the only English that should be taught to children since English exists in an expanding number of various variants.

Later, another term was introduced by [Smith \(1976\)](#) as ‘EIAL’ (English as an international auxiliary language) instead of referring to it as a foreign or second language. Afterward, a major shift was noticed in the research of English language where the focus was given to the plural varieties of Englishes such as ‘international Englishes’, ‘varieties of English’, ‘new Englishes’, ‘English languages’, and ‘world Englishes’ ([Bolton, 2012a](#)). Among these, the most popular term is ‘world Englishes’. Apart from these terms, the

localized varieties of English, new varieties of English, and non-native English have come into prominence. However, there are various approaches of research in the fields of ‘World English (es)’, ‘New English’, and ‘New varieties of English’ (Kachru et al., 2009). The notable approaches are: (1) the English Studies approach, (2) sociolinguistic approaches (sociology of language, feature-based, Kachruvian, pidgin and creole studies), (3) applied linguistics approach, (4) lexicographical approach, (5) the popularizers’ approach, (6) the critical approach, and (7) the futurology approach Kachru et al. (2009, p. 243). We briefly discuss the notable approaches as follows:

1) The English Studies approach: In the literature, a number of linguists contributed their research to the English Studies approach. In this regard, Kachru et al. (2009) in their book mentioned some of the eminent linguists, to name a few — Robert Burchfield, David Crystal, Sidney Greenbaum, Tom McArthur, Randolph Quirk, and John Wells. Each of these made an effort to address problems associated with the study and usage of English from a worldwide perspective.

2) Sociolinguistic approaches: The sociolinguistic approach includes four types of studies: a) the sociology of language, b) feature-based approaches to World English, c) Kachruvian studies d) pidgin and creole studies (Kachru et al., 2009). While describing the expansion of the English language throughout the postcolonial societies of the world, the study of Fishman (1977) is significant. These studies have offered an in-depth analysis of the expansion of English and ‘post-imperial English’ from a sociological perspective. In the feature-based approach, the linguists are concerned about the different characteristics of varieties on the basis of grammar, accent, and vocabulary. In this regard, Trudgill and Hannah (2008) played a significant role in explaining the ‘standard variations of English’, and the explanation was carried out in accordance with the differences at the level of grammar, phonetics, phonology, and vocabulary. The studies carried out by Kachru et al. (1985), Kachru (1990), and Kachru (1994) played a crucial role in describing World Englishes and its varieties. Along with other academics, the author is also credited with creating the term “World English.” Later

in this chapter, we discuss the approach proposed by Kachru (1990). In the pidgin and creole studies, the study of Görlach (1988) is of immense importance that shows the pidgin and creole varieties of the English language.

3) Applied linguistics approach: The pioneering work on World Englishes based on the ‘Applied Linguistics approach’ is the study (Halliday et al., 1964). The authors described in their work the varieties of English throughout the world, and in this regard, they stated that during the colonial era, it appeared completely evident and unchangeable that the type of English that professionals in England used was the only model that could possibly be employed in education abroad (Halliday et al., 1964). Later, the authors argued that the language English now cannot be associated with only British and Americans as in many countries like India, Pakistan, West Indies, the majority of people do not consider it to be the only accepted version of English to be followed.

4) Lexicographical approach: Kachru et al. (2009) stated that the first dictionaries of World Englishes, namely, Pickering (1816) and Bartlett (1859), were glossaries that were produced at the beginning of the nineteenth century in the United States. On the contrary, Webster (2018) felt the emergence of creating a national dictionary as it is crucial to have a language system as an independent nation. The author discussed the importance of the advent of a language of North America that will be distinct from the future language of England just like the modern European languages (Dutch, Danish, and Swedish) are different from German or from each other Webster (2018, p. 15). Similarly, in Canada, New Zealand, and Australia, various authors produced dictionaries that contain their own glossaries. Although no autonomous national dictionary of Indian glossaries and word lists has been created yet, the Indian word list and glossaries can be found in Whitworth (1885) and Yule and Burnell (1903). The importance of dictionaries in any language is immense for the recognition and establishment of that particular language. Kachru et al. (2009) mentioned that the most crucial example where the variety of English is largely accepted and institutionalized is the Macquarie

Dictionary of Australia [Delbridge and Bernard \(1981\)](#), which is accepted as ‘a national dictionary’. The editors of the Macquarie Dictionary are focusing on creating dictionaries on the other varieties of World Englishes, including the world Englishes in Asia. Due to these dictionaries, the English language would not be considered as ‘an alien language’ anymore, and these will also project the flexibility of the language ([Kachru et al., 2009](#)).

5) The popularizers’ approach: At the end of the twentieth century, an upsurge was noticed in the study of international varieties of English in Western countries. In Britain and North America, books were written by different authors describing the history of the English language. Among these studies, the study of [MacNeil et al. \(1986\)](#) and a broadcast of a nine-part BBC documentary on the history of the English language was the most popular [Kachru et al. \(2009\)](#). However, both the book and the documentary received critical responses from critics all over the world. Among the renowned popularizers of the twentieth century, the studies of [Crystal \(1988\)](#) and [Crystal \(1997\)](#) are of utmost importance. Although these works received various critical responses, the author’s focus was clear in producing the history of the English language and the position of English from a global perspective.

6) The Critical approach: In 1992, the study of [Phillipson \(1992\)](#) changed the path of the discourse of World Englishes. The author’s arguments were based on the Western liberal perspective. [Phillipson \(1992\)](#) presents a series of arguments relating to the political relationship between what the author characterizes as the “core English-speaking countries” (Britain, the USA, Canada, Australia, and New Zealand) and the “periphery-English countries” where English either has the status of a second language (e.g., Nigeria, India, Singapore) or is a foreign and “international link language” (e.g., Scandinavia, Japan) [Phillipson \(1992, p.17\)](#). The author further hints at the disproportionate relationship between the Western Anglophone powers and the developing nations where the Western countries establish their political and economic supremacy over the former colonies. It was also argued by the author that the power of West-

ern countries is accompanied by the term ‘English linguistic imperialism’, which refers to the practice of asserting and maintaining English’s supremacy through the formation and ongoing reconstruction of structural and cultural disparities between English and other languages (Phillipson, 1992). The study of Phillipson (1992) received both positive and negative comments. Apart from Phillipson (1992), another eminent work Pennycook (1994) is based on the critical perspective. Both these studies are important in describing the critical perspective of World Englishes.

7) The futurology approach: This approach refers to the future prospects of English, and in this regard, the studies of two eminent scholars, namely, Crystal (1997) and Graddol (1997) are worthy of being mentioned.

Crystal (1997) focused on the future of global English, where the author discusses various issues such as the development of New Englishes, the tendency to lean towards the mother tongue in a nation like India, and the debate about the official English movement in America. Regarding the ownership of the English language, Crystal (1997, p. 140-141) stated that “when even the largest English speaking nation, the USA, turns out to have only about 20 percent of the world’s English speakers ... it is plain that no one can now claim sole ownership”. Again, with respect to the position of English, Crystal (1997, p. 141) mentioned that “Within ten years, there will certainly be more L2 speakers than L1 speakers. Within fifty years, there could be up to 50 percent more. By that time, the only possible concept of ownership will be a global one ...”. The author also predicted the rapid spread of English and talked about the new varieties of English that are frequently referred to as ‘new Englishes’ (Crystal, 1997). The author further suggested a new dialect which is “World Standard Spoken English” (WSSE), that was claimed to be developing around the world. In this regard, it was stated by the author that in spite of having native dialects when interacting in their own nation, people would switch to WSSE while communicating with people from other countries (Crystal, 1997). In WSSE, people deliberately avoid the words and expressions that are only used in their own country and are not understood by

the people outside that country; instead, they will seek alternative forms of expression. However, the author is not sure about the prospect of WSSE in the future as it is still in the budding stage (Crystal, 1997).

The final section of Graddol (1997) speaks about 'English in the future'. Kachru et al. (2009) stated that in relation to 'world standard English', Graddol (1997) focused on some important issues, such as whether English will be disintegrated into various languages or whether American and British English will be considered to be an accurate model of English to be followed or whether the appearance of a 'world standard English' will be seen. Whereas Crystal (1997) speaks about the world standard English, Graddol (1997) opposed the idea of world standard English and predicted a 'polycentric' future for English standards outlining several studies of the economic and sociopolitical impacts of the proliferation of English.

In defining the term 'world English' and understanding its application, the study of Kachru et al. (1985) played a pioneering role in describing the concept of 'world Englishes'. According to Bolton (2012a), the term 'world Englishes' has a narrow and wide application. The narrow application of the term includes the approach of the English studies that have been carried out by Kachru et al. (1985) and a group of scholars. The wider application of the term includes the various approaches to the study of English, including the varieties that are studied throughout the world. Among these varieties, the Englishes of Britain, the USA, New Zealand, Australia, Africa, Europe, and Asia are worth mentioning. Kachru et al. (1985) modeled the use of English worldwide into three circles: 'Inner', 'outer', and 'expanding' circles. The inner circle refers to those nations where English is spoken as the 'primary language'. In this category, the countries like the USA, the UK, Canada, Australia, and New Zealand will fall. According to Kachru et al. (1985), the outer circle represents the postcolonial anglophonic context that includes large and diverse speech communities from African and Asian countries like India, Singapore, Nigeria, and Zambia. These countries are mostly multilingual, where English is used for official or co-official purposes, in legal

matters, or as a means of education. The expanding circle consists of the countries like China, Greece, Indonesia, Israel, Japan, Korea, Saudi Arabia, Taiwan, and the former Soviet Union, where English is considered an ‘international language’ or a ‘foreign language’ (Bolton, 2012b).

In this chapter of our dissertation, we discuss in detail the approaches and models proposed by Kachru (1998), Kachru (1990), and McArthur (1987). Later we discuss the dynamic model of New Englishes proposed by Schneider (1997), which is further followed by Mukherjee (2007) to explain the evolution of Indian English.

Kachru et al. (1985) mentioned that due to the rapid growth of technology and as a result of modernization, English had become an ‘additional language’ and an ‘alternative language’ in the multilingual societies of non-Western countries. According to the author, whereas the countries belonging to the inner circle shared the same cultural ideologies and political systems, the scenario is quite different in the countries belonging to the outer and expanding circles. Bolton (2012a) noted that the countries belonging to the outer and expanding circle in the model of Kachru (1990) share different cultural ideologies and political systems, and he argued that the English language achieved distinctive cultural pluralism, linguistic variety and diversity from Hinduism, Islam, Marxism, and Communism. In describing the acceptance of English as a language in the outer and expanding circle Kachru (1998) focused on the terms ‘range’ and ‘depth’. These two notions are mentioned by the author to indicate the difference between ‘genetic nativeness’ and ‘functional nativeness’. The ‘genetic nativeness’ refers to ‘the historical relationship between languages’. For example, India’s Indo-Aryan group of languages Kachru (1998, p. 92). The terms ‘range’ and ‘depth’ are associated with ‘functional nativeness’. Whereas the term ‘range’ refers to the functional use of the language in various domains such as government, law, business, family, friends, etc., the term ‘depth’ refers to the social integration of the language. The study of world Englishes is based on various issues, such as the spread and stratification of English, characteristics of the stratification, etc. While discussing the spread

and stratification of English, [Kachru \(1990\)](#) described the models of three concentric circles: the inner circle (includes American English, British English, Australian and New Zealand English), the outer circle (includes Singapore English, Nigerian English and Indian English, etc.) and the expanding circle (includes Chinese, Japanese, and Korean English). In figure 2.1, we depict the three concentric circles of World Englishes proposed by ([Kachru, 1990](#)).

[Mukherjee \(2007\)](#) stated that the term ‘World English’ is also used as ‘New Englishes’ in a broader perspective to indicate the variety of Englishes around the world. In this study, the author suggested three reasons to justify the importance of ‘New Englishes’: 1) After independence, English is widely used in former colonial territories like India, Singapore, and Nigeria for administrative, educational, and official purposes. It is also used in media and tourism as well. 2) In these countries, English is used for communication purposes, but English is regarded as the native language by a small number of people as most people use English as their second or third language for educational purposes. 3) Even though in these countries, English is not considered the first language, the emergence of various local varieties of English can be noticed, which according to Kachru’s terminology, can be stated as “institutionalized second-language varieties” ([Kachru et al., 1985](#)).

[Mukherjee \(2007\)](#) in his study mentioned two models that the linguists follow to indicate the categorization of ‘New English’. These models of categorization of ‘New English’ are important to consider and establish the new varieties of English. In this regard, [Mukherjee \(2007\)](#) mentioned two models: the wheel model by [McArthur \(1987\)](#) and the three-circle model by [Kachru et al. \(1985\)](#). The wheel model by [McArthur \(1987\)](#) has three circles. The center of the circle consists of Standard English. The second circle consists of eight regional and emerging standards of English: (1) “British and Irish Standard English” (2) “American Standard English” (3) “Canadian Standard English” (4) “Australian, New Zealand and South Pacific Standard English” (5) “Caribbean Standard English” (6) “West, East and South(ern) African Standard(izing)

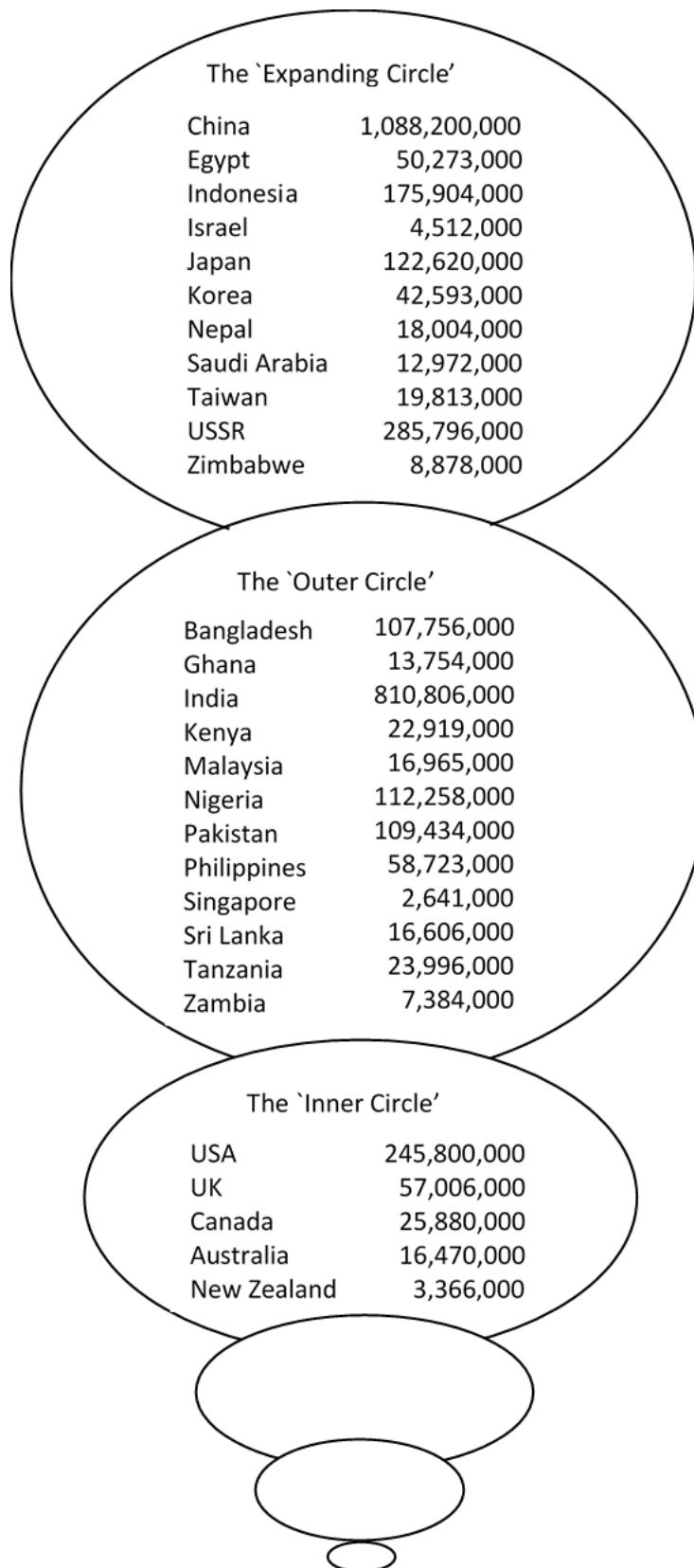


Figure 2.1: Three concentric circles of World Englishes proposed by [Kachru \(1990\)](#)

English” (7) “South Asian Standard(izing) English” and (8) “East Asian Standardizing English” [McArthur \(1987, p. 97\)](#). In this context, [Mukherjee \(2007, p. 159\)](#) pointed out the distinction among three regions—“regions with established standards, regions where standards are about to be established, and regions with an ongoing process of standardization”. The outer circle of the wheel model by [McArthur \(1987\)](#) includes the national varieties, subnational varieties, and other subvarieties. For example, Indian English, New Zealand English. Among the sub-national varieties falls the Ulster Scots and Quebec English, and the other sub-varieties include BBC English and Network Standard English ([Mukherjee, 2007](#)). In describing the significance of the wheel model by [McArthur \(1987\)](#), [Mukherjee \(2007, p. 159\)](#) pointed out that the “established native varieties and New Englishes are assigned a comparable status in terms of the process of standardization, their differentiation into subvarieties and their contribution to World Standard English”. The author also pointed out the common aspect in the wheel model by [McArthur \(1987\)](#), and the three circles model by [Kachru et al. \(1985\)](#), where both the authors focus on the concept of nativization in outer circle varieties where a foreign language is entirely indigenized and accepted. Following the concept of nativization, the indigenization of the English language in India, which can be termed as ‘Indianization’ following [Kachru \(1983\)](#), comes out to be of great importance as it has played a crucial role in bridging the gap between ‘the norm-producing inner circle and the norm-developing outer circle’ [Mukherjee \(2007, p. 160\)](#). Due to the process of nativization, apart from the native varieties of Englishes in the British Isles, North America, and South Africa, the varieties of New Englishes were also established. [Schneider \(1997\)](#) proposed a model that describes the evolution of New Englishes throughout the world from a diachronic perspective. The foundation of this model implies “that there is a shared underlying process which drives the formation of New Englishes, accounts for many similarities between them, and appears to operate whenever a language is transplanted” [Schneider \(2003, p. 241\)](#). In describing the evolution of Indian English as a variety of New Englishes, [Mukherjee \(2007\)](#) followed Schneider’s dynamic model. Here,

we discuss the analysis made by [Schneider \(1997\)](#) and [Mukherjee \(2007\)](#) and also talk about the conclusions that are derived from their respective studies.

2.6 The origin of English in India

According to [Crystal \(1997\)](#), an enormous upsurge has been noticed in the Indian subcontinent regarding the increase of English speakers in the last few decades. Even the number of English speakers in the Indian subcontinent outranks the total number of speakers in the USA and UK. It is important to note that with the use of English, distinct varieties of Asian Englishes have also emerged. This category includes Indian English along with the English that is spoken in five other countries, namely, Pakistan, Bangladesh, Nepal, Bhutan, and Sri Lanka. These varieties of English are jointly referred to as South Asian English ([Crystal, 1997](#)). After independence, among the outer circle countries, English has been used widely in various domains such as education, government, law, media, etc. [Crystal \(1997\)](#) stated that according to the last national survey, around a third of the Indian population is able to have a conversation in English, which approximately leads to a number of 350 million.

The roots of South Asian English were deeply penetrated in Britain when in 1600, with the establishment of the British East India Company, a group of merchants came for business purposes. The first trading station was established in India in 1612 at Surat. Then they gradually expanded their business in Madras, Bombay, and Calcutta. After the end of the Mughal empire, the British came into power, and along with the British, the English language was introduced into the entire subcontinent. From the year 1765, English was widely used throughout the subcontinent in official and administrative works. It was also used in education, and with the acceptance of Macaulay Minute by Lord William Bentinck in 1835, the foundation of an English education system was established in India. In 1857, the universities of Calcutta, Bombay, and Madras were founded, and English was used as the primary language ([Crystal,](#)

1997). Despite the conflicts related to the acceptance of regional languages between people belonging to various states of India, English was regarded as the prime alternative. At present, English is regarded as the ‘associate official language’ in India while Hindi is the official language Crystal (1997, p. 48). Although in the southern states of India, Hindi is considered a lingua franca, whereas English is widely used for administrative, official, and business purposes. It is also used for education, media, tourism, etc. Now, it is hard to estimate the exact number of English speakers in India as English is widely used alongside many other languages on a daily basis. Whereas some Indians use English throughout the entire day, some use English to complement their mother tongue or other languages in their multilingual repertoires (Sridhar, 2020).

As India is a multilingual country, multilingualism is deeply rooted in the history of Indian civilization. Sridhar (2013) pointed out the involvement of language in all the major events of Indian history, such as the Aryan migration; the birth of Buddhism and Jainism; the Kushana rule, the rule of Ashoka, the golden age of the Guptas, Muslim rule; the pan-Indian Bhakti movement; the Virashaiva movement of Karnataka; the rule of Shivaji, European colonization; the Independence movement; the Indian literary renaissance, and in modern times, national integration, the creation of Bangladesh, the civil war in Sri Lanka, and the Dalit movement. Speaking about multilingualism, in India, there are 122 languages (according to a census conducted by the Government of India), counting those languages only, which are spoken by more than ten thousand people Sridhar (2013, p. 3). The languages of India are from the four major language families: Indo-European family, Dravidian language family, Austro-Asiatic family, and Sino-Tibetan (or Tibeto-Burman) family. The languages such as Assami, Bengali, Gujarati, Hindi, Marathi, Konkani, Kashmiri, Marathi, Nepali, Punjabi, Oriya, Sindhi, and Urdu falls under the Indo-Aryan and Indo-Iranian subfamilies of the Indo-European family, which are mainly spoken in the northern, central, and western parts of India by 70 percent of the population. In South India, the languages like Tamil, Telegu, Kannada, and Malayalam are spoken by almost 20 percent of the population that falls

under the Dravidian language family. In the Eastern part of India, near about twenty languages of the The Austro-Asiatic family (Munda and Mon-Khmer) are spoken by about 10 percent of the population. In the Northern and Eastern border areas, 84 languages of the Sino-Tibetan (or Tibeto-Burman) family are spoken by one percent of the population (Sridhar, 2013). Gupta et al. (1995) stated that in India, 22 languages relish the special status of 'scheduled languages' as these languages are mentioned in the Eighth Schedule of the Indian Constitution. These scheduled languages include Bengali, Assami, Bodo, Dogri, Gujarati, Kashmiri, Hindi, Kannada, Konkani, Malayalam, Manipuri, Marathi, Maithili, Nepali, Oriya, Punjabi, Sanskrit, Santhali, Sindhi, Tamil, Telugu, and Urdu. This list of scheduled languages involves Sanskrit, the pan-Indian classical language that is left with few native speakers, and Urdu, which is regarded as the identity language of the Muslims and is used by 15 percent of the population. It also contains literary languages such as Bengali, Hindi, Tamil, and Kannada, along with some tribal languages, for example, Santhali and Bodo. However, these scheduled languages are commonly known as regional languages in India. There are 29 states and 7 union territories in India, and 90 percent of the population speaks one of the above-mentioned languages (Sridhar, 2020). In India, more than ten million people use the scheduled languages, but in a way, they also form minorities nationally as Hindi is the most extensively spoken language throughout the country, and the regional languages are somewhat restricted to the respective states or regions. In this regard, it is worthy of mentioning that the Bengali language is an exception as it is used for interaction and communication in most of the North-eastern states of India. There is no monolingual state in India, and a wide range of diversity can be seen among the languages. Although multilingualism in India is considered to be a part of formal schooling, generally, it is acquired informally from family, neighbourhood, and workplace.

In this section, we discuss the emergence of Indian English and how it gained its status around the world along with the New Englishes. First, we discuss the dynamic model of the evolution of the New Englishes proposed by Schneider (2003), and next,

we talk about the model proposed by Mukherjee (2007) who followed Schneider (2003)'s model to describe the historical development of Indian English and its present status.

Schneider (2003) proposed a dynamic model of the evolution of the New Englishes. A number of researches have been carried out focusing on the features of the individual varieties of New Englishes. Although some authors have identified the similarities between the varieties of New Englishes, these similarities are noticed within a particular region, for example, south or Southeast Asia. Apart from the similarities, there are differences that are caused by the process of colonization, various historical incidents, linguistic variation, cultural and regional factors, and thinking processes. While proposing the dynamic model, Schneider (2003) discussed that there is a common underlying process that creates New Englishes, explains many of their similarities, and seems to be active whenever a language is transplanted. The author further characterized this process by two interrelated factors: 1) changing identity construction and 2) changing interactions between settlers (STL) and the indigenous population (IDG). The first factor indicates the linguistic change that was noticed in the settlers who consider themselves "as an extension of the 'us' of their country of origin as against the 'other' of the indigenous population of their country of destination" Schneider (2003, p. 242). But later with the course of time, the formerly inhabited nation transformed into 'other', and the newly developed population was regarded as 'us'. The second factor refers to the socio-political and communicative relationship between the colonizers and the colonized. According to Schneider (2003, p. 242) "to a considerable extent the histories of New Englishes can be viewed as processes of convergence between these two groups, despite all the initial and persistent differences between them". Further, Schneider (2003) claimed that these two factors are responsible for an evolutionary pattern in the formation of New Englishes, which is applicable to the varieties of English worldwide. This evolutionary pattern is composed of five successive stages: (1) foundation, (2) exonormative stabilization, (3) nativization, (4) endonormative stabilization, and (5) differentiation. These five stages are discussed as follows:

In the foundation phase, because of the English-speaking settlers, English was the language that was used on a daily basis in a territory where the speakers were not English. The settlers continued to follow the language norms of their home country Britain and the communication was taking place on two levels: a) internal communication between the English-speaking settlers and b) communication between the settlers and the indigenous population (IDG). The communication with the IDG strand is restricted to trade and political or military negotiations (Schneider, 2003). During this time, a group of bilingual speakers developed who used to play the role of interpreters and take a major role in communication. Though at this phase, the two strands (STL and IDG) remain more or less separated (Mukherjee, 2007).

In the exonormative stabilization phase, the STL strand still considered themselves as outposts of their former home country and still followed the British culture in the foreign land. The STL strand at first sticks to the English that they used to speak in their homeland. But later, they started to adopt the local vocabulary and thus gradually adhered to the localized form of English. During this time, the names of local animals and plants were included in the English vocabulary. Also, the local English community in India developed that includes the STL strand who were aware of their British origin; however, a change can be noticed in their identity, which according to the author, was “genuinely British no doubt, but seasoned with the additional flavor of the extraterritorial experience” Schneider (2003, p. 246). On the other hand, the IDG strand started to learn English as a foreign language. Through education and communication in business and trade, bilingualism gradually spread among the IDG strand. For them, learning and mastering the English language paved the path for better opportunities in educational as well as commercial fields, and it also helped them to be a part of upward social mobility. With this, a positive attitude towards learning and using the English language emerged among the indigenous population. The English knowing people of the community became popular for their ability to communicate with the English and the Europeans, even after being a part of the community. Also, during

this time, despite being two separate identities, extensive communication continued to take place between the STL and the IDG strand (Schneider, 2003).

In the evolution of New Englishes, the third phase called nativization is hugely important as in this phase both the STL and IDG strands developed a new identity. Both STL and IDG strands felt the need to become close to each other as both groups started to consider themselves permanent residents of the territory. During this time a new local identity was formed which was not reflected in every sphere of ‘linguistic, social, and political reality’ Mukherjee (2007, p. 162). A local variety of English appeared at this time that contains new forms and structures in grammar and vocabulary. On the other hand, a conservative attitude also appeared among the STL strand that is referred to as “complaint tradition”. Schneider (2003, p. 248) stated that “conservative language observers typically claim that linguistic usage keeps deteriorating, that in the new country ‘corrupt’ usage can be heard which, however, should be avoided”. A difference can be noticed among the STL strand where the more conservative users still followed the ‘traditional exonormative (British) standard’, while the progressive language users started to use more the ‘locally emerging norms’ (Schneider, 2003).

The fourth phase is called the endonormative stabilization where with the completion of the nativization process, locally rooted self-confidence arises among the indigenous strand. Due to this self-confidence, the tendency to lean toward Britain became no longer beneficial. For the transition from Phase 3 to Phase 4, an essential requirement was the political independence of the colony. According to Schneider (2003, p. 250), in the post-independence period, ‘some exceptional quasi-catastrophic political event’ or an “Event X” triggers Phase 4, and in this phase, the local norms were accepted and a change has been carried out on a lexico-grammatical level. During this time, the use of English for writing creative fiction increased to a large extent which results in the creation of “New Literatures in English” or “Commonwealth Literatures” (Mukherjee, 2007). In describing the characteristics of the local norms it is stated in Mukherjee (2007, p. 163) that “In phase 4, the homogeneity of the local norms, serving

as positively evaluated carriers of a local identity, tends to be emphasized, which is usually mirrored in nation-based labels for the new variety (“X English”-e.g., Indian English)”.

In the fifth phase called differentiation, instead of creating a new national identity, the emphasis was on creating new subnational group identities within the new, stable nation (Mukherjee, 2007). During this time, various English dialects came to use due to the local linguistic norms. Schneider (2003, p. 254) also stated further that “phase 5 does not entail monolingualism in English at all” and “it is possible for varieties of English to coexist with other, mostly indigenous languages, with all of these fulfilling identity-marking functions.” Mukherjee (2007) stated that in order to be significant power and create an identity at national and sub-national levels, the English language must be extensively used in a monolingual and multi-lingual setting. This is true both in terms of the range of functions it performs and in terms of how deeply it is ingrained in society. The evolutionary pattern proposed by Schneider (2003) is immensely important in the formation of the New Englishes around the world. Although regarding the phases it is often argued that the phases are often not distinguished from one another, the model clearly upholds its claims: 1) it shows a uniform pattern that lays the foundation of the formation of New Englishes, 2) every country around the world where New English evolves will conclude at some point in phase 5 (Schneider, 2003). However, even after being the largest ESL variety, Indian English is not included in the study of Schneider (2003). While it is argued that only 5 percent of the Indian population uses English fluently and regularly in India (Tully, 1997), Sheorey and Nayar (2002, p. 14) stated that “ even that small percentage translates into about 50 million users of the language, making India the third largest English-using country after the United States and Great Britain”. Later, Mukherjee (2007) followed Schneider (2003)’s evolutionary model of New Englishes and described the evolution and development of Indian English following the five phases.

1) **The Foundation Phase:** The foundation phase marks the introduction of the

English language in India. With the arrival of the first Englishman named Father Thomas Stephens in the country in 1579, the arrival of the language also took place. Along with him, the evolution of 'Anglo-Indian Literature' started (Ward and Waller, 1916). In 1600, the East India Company came to India for trade, and along with that came various English merchants. For trading purposes, these English merchants used to communicate with the Mughal Emperors and also with the common people. In the evolution of the English language in India, missionaries played an important role. They attracted the attention of the common people of India. During this time, the British army also came to power, and they attracted the attention of the Indian soldiers. However, it must be noted that the English colonizers always considered themselves as part of their homeland and continued to carry the identity as British (Mukherjee, 2007).

2) **Exonormative Stabilization phase:** With the decline of the last Mughal emperor named Aurungzeb, India was already involved in many battles with the British, the French, and the Muslim leaders in the North and South of India (Mukherjee, 2007). This opened up the path for the British to get more involved in the rivalries of the country. Using this chance, the British also established their footsteps firmly in the coastal areas, mainly Mumbai and Bengal. The dawn of the British Empire in India started with the victory in the Battle of Plassey in 1757, which was led by Robert Clive. During this time, the Mughal Emperors were still ruling in Delhi. The victory of the British in the Battle of Plassey can be marked as the beginning of the British Empire. After the victory, the Mughal Emperors were forced to grant the political and administrative authority called the Diwani to the East India Company for Bihar and Bengal (at that time, East and West Bengal were jointly known as Bengal). Mukherjee (2007) stated that in 1773 with the Regulating Act, the East India Company turned into a British administrative body, and with the East India Bill of 1784, the control of the East India Company was shifted from the British Parliament to Her Majesty's government. However, it was argued by Mukherjee (2007) that the exonormative sta-

bilization phase of Indian English started in 1760. During this time, both the STL and IDG strands of India understood that the British were not leaving the country soon, and as the new power rose, the language of the rulers also gained importance. During the nineteenth century, a large increase was seen in the British population in India, and the number of missionaries also increased throughout the country, which helped in the spread of the English language among the Indians. Similarly, an upsurge in the number of Indian soldiers in the British-Indian Army was also seen. During this phase, some regularly used Indian words, mainly the name of some Indian items (for example, curry, bamboo, mango, veranda, etc.), were included in the English language (Mukherjee, 2007). Although in this phase, the English language got the touch of Indianism, the standards and norms of English still remained British in India. In the early nineteenth century, an influential group of people was attracted to Western education, science, and culture. Mukherjee (2007) in his study mentioned that the Orientalists were in favor of an Indian education system that centers around Indian languages, culture, and literature. On the contrary, the Anglicists, represented by Thomas Macauley, with his Minute on Indian Education (1835) forced to establish an English medium education system in India as they assumed English culture was way superior to Indian culture. As a reason, the Anglicists argued that an English education system was necessary to get loyal Indian civil servants. Thus, the establishment of the English medium school system started in India, and obviously, English became the prime language of instruction in secondary schools and also in the very first universities, for example, Bombay (now known as Mumbai), Calcutta (presently known as Kolkata) and Madras (presently known as Chennai).

3) **Nativization phase:** The nativization phase in India started with Macauley's (1835) Minute (a legislative act that advanced the spread of Western education policy in India). Thomas Babington Macaulay firmly believed that in India, the Western education system must be implemented as it will be a medium of control, along with the moral responsibility to educate the native people of the country with Western

literature and education. The proclamation of Macaulay was accepted by the then Governor General of India, Lord William Bentinck. The Indian people who adopted the English education system performed the role of mediators who filled the gap between the British and the Indian culture. During this time, English was widely used in administrative works, the military, and the jurisdiction of India and also in the higher education system as this would serve the purpose of the British to get people for lower ranks in Indian Civil Service. With the beginning of the English Medium education in India, English literature and culture also entered India. In the long run, it showed an adverse effect on the British as with English literature and culture, Western ideas like democracy, enlightenment, and self-determination also entered the hearts and minds of the Indian people. Regarding this, Jawaharlal Nehru, the first Prime Minister of India who also actively participated in the freedom movement, stated that with English education, an expansion of the Indian perspective was observed, which includes acclamation of English literature and English institutions. Along with that, a rebellious tendency towards some Indian aspects and customs was noticed among the Indians, which eventually upsurged the need for political reform. A new class, inspired by Western ideas and practices, evolved that contained the English-educated Indians who were isolated from the majority of the population. The initiation of the British Indian Association in Calcutta was in 1892. This was one of the precursors to the Indian National Congress, but it would take another generation for the Congress to begin operations in 1885 (Nehru, 2008). Although the British thought it would be beneficial for them and the English language would act as a communicative device to interact within multilingual India, in reality, the Indian intellectuals from the entire subcontinent used English to revolt against the British rule (Mukherjee, 2007). On the contrary, the British felt at home in India as the whole subcontinent became an increasingly Anglophone territory. Mukherjee (2007) mentioned that in the mid-nineteenth century, there was an emergence of a local English-based identity that consisted of both the British and the Indian locals, and this was the outcome of the intertwinement

of both the STL and the IDG strands. In 1857, there was a massive revolt led by the Indian soldiers in Meerut in order to drive out the British from the country. This great revolt is known as Indian Mutiny or Sepoy Mutiny. The British also revolted with the help of the other Indian troops who did not participate with the Indian soldiers, and in the end, the British won, and they strongly reestablished their rule in India. According to Mukherjee (2007), in 1877, the successful establishment of British rule in India was concluded with the declaration of Queen Victoria as Empress of India, and at the same time, a virtually omnipotent Viceroy was also appointed who would represent the British Crown in India and reign as an absolute monarch. These events paved the path more widely for English to be the language of power and dominance in India. During this time, the upsurge of the British population could be seen in India, and the British colonial power became an inseparable part of Indian politics and identity.

From the mid-nineteenth century, English was the language that was necessary for higher education and for Indian civil service. During this time, English was widely used by the IDG strand, and a variety of English was gradually growing that included not only lexical borrowings but also phraseological and grammatical innovations. The changes include some phonological changes, for example, monophthongization of diphthongs (Mukherjee, 2007). The growth of ‘educated’ Indian English was expanded in the late nineteenth century and could be termed according to Kachruvian terminology “Standard South Asian English” (Kachru, 1986). At this point, the language of the STL strand had become a medium for communication and interaction between the IDG strand. After India gained independence from the British on 15th August 1947, English was used by the freedom fighters in the Indian National Congress as a useful pan-Indian link language. Again, the position of English in the country solidified when the first Prime Minister of independent India, Jawaharlal Nehru, delivered his speech in English (Mukherjee, 2007). After independence, the process of nativization of English in India still continued. Ironically, the Constitution of the Republic of India was

also written in English that was passed by the Constituent Assembly in 1949 and came into effect in 1950. Basu et al. (1999) mentioned that it is only the English version of the constitution that is legally binding even today, even though English was not listed among the eighteen official national and regional languages in the Indian Constitution. After some years of independence, the process of nativization of English in India was thought to have stopped as there were provisions in the constitution to replace English with Hindi, which was the mother tongue of 35 percent of Indians. But surprisingly, in post-independence India, English was announced as the second official language of India, rejecting the very thought of going through the process of ‘denativization’. Thus, the English language in India successfully overcome the process of nativization and proceeded to enter phase 4, which is endonormative stabilization.

4) **Endonormative Stabilization phase:** This phase usually describes present-day Indian English. Mukherjee (2007) claimed that in order to enter the phase of endonormative stabilization, a language needs an inner agreement in the speech community to assert the status and usefulness of that language without any external involvement. In the constitution, it was predicted that English alongside Hindi would remain the pan-Indian link language till fifteen years after the independence. During the 1960s, various conflicts arose across the country, especially in the northern and southern parts of India. In the northern parts of India, Hindi was considered the only national language, whereas, in the southern parts, people rejected Hindi as the national language as Hindi was a non-native language to them. In 1965, the language riots reached their peak mainly in the southern state of Tamil Nadu, that led the political parties to rethink their language policy and to confirm the continuous use of the English language in India. In 1963, the Official Language Act was passed, which was amended in 1967 and confirmed the use of English alongside Hindi for official purposes. In 1976, official language rules were formulated to specify the various official communication situations at federal and state levels in which Hindi and/or English were to be used (MIB, 2001). Regarding English language teaching in India, a compromise was

made between the supporters of English as an official language and the supporters of Hindi as the only official language and a three-language formula was decided. This formula asserted that three languages, Hindi, English, and a regional language, would be taught in every state (Biswas, 2004). By keeping aside all the major problems regarding language priority, the three-language formula has been at the heart of language policy in the education system of India over the past four decades (Krishnaswami and Sriraman, 1995). Kachru (1986) stated that English was finally accepted as an “integral part of India’s linguistic repertoire” and it is the joint result of the language riots of the 1960s and the legislative reaction.

Despite the fact that the official objective of the federal language policy asserted the propagation of Hindi as the only national language, the extensive use of English throughout the nation secured its status of being regarded as the second official language of the union, further established various labels, for example, “associate additional language” and “associate official language” (Mehrotra and Mugler, 2000). In addition, the English language was regarded as the only official language at the federal level and one of the primary official languages of the four states and union territories, namely, Chandigarh, Meghalaya, Mizoram, and Pondicherry (Mukherjee, 2007). Thus, the English language in India has gone through the endonormative stabilization phase. Mehrotra and Mugler (2000) pointed out that English has been retained in a wide range of communication situations, including administration and politics, education and academia, and press and book publications, and it has been increasingly used as a pan-Indian link language. Mukherjee (2007) argued that presently Indian English could be considered as a “phase 4 variety” in Schneider (2003)’s model as it has passed the process of nativization and stabilized itself endonormatively. Further, Mukherjee (2007) argued that although Indian English is endonormatively stabilized, it has not gone through the process of differentiation as it is not entirely homogeneous across the subcontinent regarding the aspects of pronunciation, lexis, and grammar. Also, the author pointed out that the phases in Schneider (2003)’s dynamic model could be seen

as shading into each other, it is not surprising that the features of phase 3 still exist in phase 4 of Indian English. There are various observations that indicate the traces of nativization which still remain in phase 4 of Indian English. According to Mukherjee (2007), such observations include the tendency in formal written English to cling to lexicogrammatical standards set by native speakers and “linguistic schizophrenia”. Kachru (1986) labeled the term “linguistic schizophrenia” that refers to the tendency of South Asian speakers who rejects the local variety of English and simultaneously accepts English as an integral part of their linguistic repertoire. Mukherjee (2007) also argued that in Indian English, some signs of the emergence of dialectal differentiation could be seen, which was similar to phase 5 of Schneider’s dynamic model. Here, an example of dialectal differentiation by Baldauf (2004) would provide a clear picture: in Mumbai, people use the word ‘stadium’ to refer to men with a bald spot with hair all around (though it was only used in an informal context).

So it is evident that both the studies of Schneider (2003) and Mukherjee (2007) have come to different conclusions. Whereas according to the evolutionary model by Schneider (2003), Indian English is at phase 3, or nativization phase, Mukherjee (2007) argued that Indian English is at phase 4, or endonormative stabilization phase. There is no doubt among scholars that Indian English meets the characteristics of phase 3, i.e., nativization phase (Sailaja, 2012). However, scholars have a contradictory opinion about placing it in phase 4 due to the homogeneity issue. Sailaja (2012) argued that the issue of norms in Indian English is still unresolved and urged for a detailed investigation to resolve ambiguity.

2.7 Characteristics of Indian English

In this section, we discuss in brief the features of Indian English. It is important to mention that as British English and American English can be marked by their own recognizable accents, Indian English also carries a distinct Indian accent. Earlier,

Kachru (1994) stated that on the basis of language families, divisions of Indian English are made, for example, Indian English of the Indo-Aryan language background and Dravidian language background. Despite the variation regarding the accent, there are some common features throughout the nation that helps to identify the particular variety of Indian English. Here, we discuss the phonetic and phonological features, the lexical and morphological features, and the syntactic features of Indian English.

2.7.1 Phonetics and Phonology

There are some common phonetic and phonological features in Indian English. Sailaja (2012) mentioned the distinct Indian accent that is one of the most important characteristics of Indian English which is acquired from the L1 speech patterns of the speakers. However, in India, there are different language families, and the accent varies according to the different language backgrounds (such as Indo-Aryan language background and Dravidian language background) of the speakers. Despite the difference regarding the accent, there are some common features that help to identify the variety of Indian English. Regarding the first classifications of the segmental features of Indian English, both the works of Masica (1972) and Bansal (1976) are considered to be of utmost importance. Henceforth, various researchers such as Pingali (2009), McArthur (2002), Trudgill and Hannah (2008), Pandey (1994) etc. have also explored Indian English. The common features that are commonly identified in Indian English are retroflex sounds /ʈ, ɖ/. These sounds are the initial sounds of the words like ‘ton’ and ‘dozen’. In some cases, the alveolar sounds /t, d/ can also be present with the retroflex sounds in the speech of some speakers. These features vary from speaker to speaker, although the variation depends on the context (Pingali, 2009). In the native varieties of English, fricative sounds /θ, ð/ are used in place of the dental plosives /t̪, d̪/. For example, in words like ‘thanks’ and ‘these’, the plosives are in the initial position. Some speakers also use labio-dental approximant /v/ instead of both /v/ and /w/ (Wiltshire and Harnsberger 2006, Pingali 2009). In Indian English, in the place of diphthongs,

long vowels are used, and instead of rounded back vowels, low back vowels are used. For example, ‘may’ [mei] as [me:], and the word ‘coffee’ is pronounced as [kafi]. Also, affricates are used for fricatives, for example, [p] for [f], [s] or [dz] for [z], as can be seen in the word ‘zero’ [siro] or [jiro]. In some South Indian varieties of English [ʃ] is used in the words like ‘treasure’ and ‘pleasure’ (Sridhar, 2020).

There are few works that describe word stress, intonation, and rhythm of Indian English, among which the studies of Pandey (1994) and Gargesh (2008) discuss stress placement in Indian English. The stress is usually placed on the heavy syllable, as in the case of ‘moñsoon’. However, if there is no heavy syllable available, then the stress is usually placed on the first syllable as in ‘dévelop’(Sailaja, 2012). Chaudhary (1989) investigates two different stress patterns in Indian English that are based on the speakers of particular language families where mother tongue influence is also considered to be important. For example, in the case of the Dravidian language speakers, it is échemistry whereas, in the case of the Indo-Aryan group of speakers, it is chémistry. In Indian English whether the rhythm is syllable-timed or not is a matter of argument. However, Babu (1971) pointed out that it is not syllable-timed or stress-timed. In this regard, the study of Pingali (2009) stated that in Indian English, it is not syllable-timed rhythm as in this case equal importance is given in the articulation of each syllable of a sentence. Regarding intonation, Pingali (2009) pointed out that the falling tone is usually preferred first and then the rising tone.

2.7.2 Morphology and vocabulary

Morphology: According to Sailaja (2012), though there are a number of studies that investigate Indian English vocabulary, no notable work has been carried out on word formation and morphology within the sphere of Indian English. However, there are some words that are the result of compounding which also includes hybrid compounds. In hybrid compounds, one element is from English, and the other is from any of the Indian languages. For example, the word ‘lathi-charge’ means the action of beating

with a stick or lathi (Hindi) (Sridhar, 2020). Another example is the word ‘burning ghat’ which refers to a platform near the riverside where the Hindus usually cremate the dead bodies. Compounding is also seen with the Indian language affixes. For example, the suffix ‘wala’ is often used with various words from Indian languages: ‘police wala’ (policeman), ‘chai wala’ (tea seller), ‘paper wala’ (newspaper delivery man).

The affixation process in Indian languages includes identifying a person by the place of their origin. For instance, the words like ‘Naxalite’ (a person belonging to the naxalbari region), ‘delhiite’ (a person who belongs to Delhi or is an inhabitant of Delhi). Presently, the word ‘naxalite’ doesn’t only refer to a person who belongs to the naxalbari region rather it also refers to the leftists who involves in armed rebellion against the state (Sailaja, 2012). Pingali (2009) mentioned various suffixes like ‘ese’ (from English) as used in the word ‘Assamese’, the suffix ‘i’ (from Hindi) used in the word ‘Hyderabadi’, Madrasi, etc. Trudgill and Hannah (2008) mentioned about pluralization of the uncountable nouns or mass nouns such as ‘aircrafts’, ‘furnitures’ etc. There are some examples of participial compounds such as ‘convent educated’ (refers to a person who is educated in a convent school), ‘foreign-returned’ (refers to a person who has returned from abroad) (Sridhar, 2020, p. 13).

Vocabulary: D’souza (2001) claimed that in Indian languages English expressions are widely used, and it is deeply rooted in the Indian ethos. Talking about vocabulary, according to Sailaja (2012), there are several Indian words that are added to the English vocabulary and are also used in different varieties of English. For example, the word ‘avatar’ means virtual projection, ‘chai’ means tea, ‘guru’ used to mean teacher or master, ‘mantra’ means formulaic refrain, ‘chutney’ means mixed relish, and ‘samosa’ used for vegetable fritter. Again, the word ‘shampoo’ is also derived from the Indian subcontinent which means ‘to massage,’ (Rao, 1954). There are some words that are particularly used in Indian English, and they are not used in the native varieties of English(es). For example, the words like ‘clever’ (which means intelligent), ‘drama’ (means play), and ‘stir’ (means agitation) (Sailaja, 2012). The other examples include

‘desi’ (means local or native), ‘aam aadmi’ (means common man), ‘accha’ (used for OK), ‘crore’ (used to mean 10 million), ‘babu’ (means bureaucrat), ‘yaar’ (to refer to someone who is a close friend), ‘filmi’ (to mean melodramatic), ‘saheb’ (to refer to an officer). There are also some words that are particularly restricted to the regional sub-varieties of Indian English. For example, specifically in North India the word ‘chowkidar’ is used to refer to a security guard, and ‘memsaheb’ is used to mean madam or a lady boss. In Bengal, the word ‘bhadralok’ is used to refer to cultured people, and in South India, the word ‘outhouse’ is used to mean a small adjoining house on the same property (Sridhar, 2020, p. 14).

2.7.3 Syntax

According to Sailaja (2012), the linguistic form that has the least variation across the varieties of English is syntax. Parasher (1983) claimed after studying some Indian samples given to British, American, and Indian speakers of English that there is no remarkable difference between Indian English and the other varieties of English. Nevertheless, there are some studies that are carried out to identify the syntactic features of Indian English. The study of Agnihotri et al. (1988) explored some Indian features while describing how the college students of Delhi use tenses. Sridhar (2020) stated that among the various syntactic features of Indian English, the influence of the mother tongue plays a crucial role. For instance, the use of articles in Indian English is different. In Indian English, indefinite articles are not used in some cases, whereas they are used in the native varieties. Also, the opposite can be noticed in some cases where indefinite articles are used, whereas the usage of those is not seen in the native varieties. Another syntactic feature of Indian English asserts the intransitive use of transitive verbs such as ‘No problem, we will adjust’, ‘We enjoyed very much’ (Sridhar, 2020, p. 16). In Indian English, there is an important syntactic feature called inversion. In both Wh-questions and yes-no questions, the lack of inversion regarding the subject and auxiliary is noticed, and this feature is most common in South Asian languages.

For example, ‘And when this puja is to be started?’, ‘Then, where to do this puja?’. Also, in embedded questions, the inversion of the subject and auxiliary is not direct, for example, ‘When you are going to London?’ (Sridhar, 2020, p. 16). Sridhar (2020) investigated the word order in Indian English. Due to the substratum influence, in the South Asian languages, the word order is subject-object-verb (SOV), whereas in English the word order is subject-verb-object (SVO). As a result of this feature, topicalization is widely used by Indian speakers. For example, ‘This I like’. Here the object phrase is moved to the left of the verb. The other syntactic features of Indian English include the use of invariant tags in tag questions, for example, ‘You aren’t going home, isn’t it?’, ‘Those hooligans are destroying public property, isn’t it?’, the enormous use of participial phrases and clauses as in ‘Having climbed the hill, and having looked into the cave, they decided to rest for a while’, the use of redundant prepositions (the use of two prepositions simultaneously in a row) as in ‘discuss about’, ‘mention about’, and the use of plural marker with collective nouns such as ‘furnitures’, ‘luggages’ etc. The use of the words ‘also’ and ‘only’ is different in Indian English as they are often placed at the end of the sentence, for example ‘But then as now, India was like that only’ (Sridhar, 2020, p. 17).

In South Asian languages, often Indian English speakers answer the questions based on the truth of the underlying presupposition where the form of the question is not what mattered. For example, if someone is asked ‘Sudhanva, I hope you are not hurt badly?’, the answer to this question would be ‘Yes, uncle, I am not hurt badly’. The difference can also be seen regarding spelling as in Indian English spelling is followed as they are used in British English such as ‘colour’, ‘realise’, ‘sceptical’ etc (Sridhar, 2020, p. 17). Another feature pointed out by (Sridhar, 2020, p. 18) is the change of Anglicized spellings of the Indian places (named during the British rule) according to the correct Indian pronunciation. For instance, ‘Kolkata’ for Calcutta, ‘Mumbai’ for Bombay, ‘Bengaluru’ for Bangalore, ‘Puducheri’ for Pondicherry, etc.

2.8 Conclusion

This chapter concentrates on the emergence of the English language in Britain as well as in America. Also, it describes World Englishes and the spread of the English language in the Indian subcontinent during British rule and how it evolved during the ages. The features and characteristics of British, American, and Indian English are also demonstrated here in brief. We mentioned some of the basic morphological, syntactic, phonetic, phonological, and lexical differences between British and American English. We also investigated these features in the sphere of Indian English. Presently, the English language is used in most countries throughout the world either as their first or second language. The rapid growth of English has secured its position as a language of international communication, and it is also considered a first language by more than 400 million people all over the world. The fact cannot be denied that at present English has become an integral part of our daily lives. The purpose of this chapter is to understand the advent of the English language in the three nations (UK, India, and USA) and to understand the cultural background of the writers (in our case, the commenters, who comment and reply on specific news). In the next chapter, we discuss metadiscourse markers in detail.

Chapter 3

Metadiscourse

3.1 Definition

Metadiscourse refers to the ways in which writers interact with their readers. It helps to organize the text by engaging the audience and by giving a clear glimpse of the writer's attitude. In recent years, metadiscourse has been extensively studied and used in pragmatics, discourse analysis, academic writing, and language teaching. Metadiscourse is said to be "discourse about the evolving discourse" (Ädel, 2006, p. 2) or "writing about writing" (Williams and Nadel, 1989). It helps the readers to understand the exact direction of the text by giving some hints through which the readers can easily understand the progression of the plot. It is important for both verbal and literary communication. According to (Hyland, 2005a, p. 43), metadiscourse is "the cover term for the self-reflective expressions used to negotiate interactional meanings in a text, assist the writer (or speaker) to express a viewpoint and engage with readers as members of a particular community." Crismore (1983, p. 2) stated that "Metadiscourse is, simply, an author's discoursing about the discourse; it is the author's intrusion into the discourse, either explicitly or non-explicitly, to direct the reader rather than inform. Metadiscourse is the directives given to readers so they will understand what is said and meant in the primary discourse and know how to 'take' the author." Later, Biri (2018,

p. 154) claimed that “metadiscourse is a pragmatic framework for describing explicit, linguistic devices for organizing the text flow, engaging the audience, and expressing the author’s stance.”

Metadiscourse is essential for both the writer and the reader as it creates a connection between them by allowing the writer to guide the readers throughout the text and by allowing the readers to participate in the text. It upholds the ways through which the writers unveil their communicative intentions within the discourse. The most interesting thing about the use of metadiscourse is that it is not attached as additional information in the text like a footnote; rather, it is associated within the text in the form of words and phrases. The metadiscourse resources can occur as a single word or phrase, or they can occur as a whole sentence, even as a whole paragraph of a written text. [Boshrabadi et al. \(2014\)](#) mentioned in their study that metadiscourse is important for both the teachers and the foreign language learners in order to understand the organization of the text. According to [Crismore et al. \(1993\)](#), the term metadiscourse refers to non-propositional features of discourse by the help of which prose is organized as a coherent text, and a writer’s personality, integrity, reader sensitivity, and connection to the content are revealed.

However, the term ‘metadiscourse’ is not always used in a similar way as, according to some analysts, for example, [Mauranen \(1993\)](#) and [Valero-Garcés \(1996\)](#), the range of metadiscourse is contracted to the ways that organize the text. Contrarily there are various analysts and researchers who explained and investigated metadiscourse in various genres and disciplines. While [Schiffrin \(1980\)](#) discussed the importance of metadiscourse in casual conversation, [Crismore and Farnsworth \(1990\)](#) talk about the importance of metadiscourse in science popularizations, [Swales \(1990\)](#) explained the importance of metadiscourse in postgraduate dissertations. [Hyland \(1996\)](#) showed the importance of metadiscourse in science research articles. The use of metadiscourse is also seen in the works of [Cheng and Steffensen \(1996\)](#) and [Intaraprawat and Steffensen \(1995\)](#), which investigated the use of metadiscourse in ESL essays and native-speaker

student writing. Thus metadiscourse not only organizes a discourse but also helps to ensure the writer's position, and it builds a relationship with the audience with the help of a communicative process. (Hyland, 1998b, p. 438) pointed out that metadiscourse cannot be used by the authors according to their will as 'it is integral to the contexts in which it occurs and is intimately linked to the norms and expectations of particular cultural and professional communities'. Regarding the effective use of metadiscourse, (Hyland, 1998b, p. 438) further stated that it depends on 'a rhetorical context and the writer's observation of appropriate interpersonal and intertextual relationships.' It is claimed by Hyland (1998b) that metadiscourse is the ways through which we negotiate with the readers or listeners while speaking or writing about how the decisions are affecting them. The metadiscourse markers not only help the writer to involve in the text but also engage the readers. Talking about the importance of metadiscourse, it is worth mentioning that without the metadiscourse markers, it will be hard to follow a text or passage, and it will be less interesting also.

In this chapter of our doctoral dissertation, we discuss metadiscourse in detail and how it helps in analyzing and organizing the text. There are various metadiscourse markers through which the writers communicate with their readers. Here, we describe some of the classifications and categorizations of metadiscourse.

3.2 Classification

In this section, we first describe the studies on academic writing, next we proceed to mention the studies on newspaper articles, editorials, news reports, and opinion columns. Then, we present the metadiscourse studies based on the writings of the EFL learners. We conclude this section by mentioning the studies carried out on blogs, posts, and news comments. Previous studies prove that there are different categories of metadiscourse. Thompson (2001, p. 58) categorized the process of interaction between the reader and the writer into two types: interactive and interactional, where

the author stated that “interactive resources help to guide the reader through the text, while interactional resources involve the reader collaboratively in the development of the text”. The author considered the concept of the ‘reader-in-the-text’ in his study to analyze the role of interaction in the written text. In 1981, Williams (1981) categorized metadiscourse into three common types: hedges and emphatics, sequencers and topicalizers, and narrators and attributors. Later, Williams reclassified metadiscourse into three general types: advance organizers, connectives, and interpersonal discourse. In 1975, Meyer (1975) used another classification system for signaling. The author identified four major types that include: 1) the specification of the structure of relations in the content structure, 2) prospectively revealed information abstracted from content occurring later in the text, 3) summary statements, and 4) pointer words.

In 1983, Crismore (1983) presented a metadiscourse model that is based on both Williams (1981) and Meyer (1975)’s classification of metadiscourse with some modifications. The author classified metadiscourse into two general categories: informational and attitudinal. The informational metadiscourse is further classified into four types: 1) global goal statements, 2) global preliminary statements about content and structure, 3) global review statements about content and structure, and 4) local shifts to the topic. The author describes the global goal statement as the goal, the global preliminary statement as pre-plans, the global review statements as post-plans, and the local shift to the topic as topicalizers. Similarly, the attitudinal metadiscourse includes the following sub-types: 1) saliency, 2) emphatics, 3) hedges 4) evaluative. According to the author, saliency refers to the importance of the idea, emphatics refer to the degree of certainty of assertion, hedges refer to the degree of uncertainty, and evaluative refers to the attitude towards a fact or idea.

Another model of metadiscourse was presented by Crismore et al. (1993), where the author divided metadiscourse into two categories, namely, ‘textual’ and ‘interpersonal’. The textual metadiscourse is divided into two sub-categories: text markers and interpretive markers. The interpersonal metadiscourse is further divided into five sub-

categories: hedges, certainty markers, attributors, attitude markers, and commentary. The textual markers help to organize the discourse, and the interpretive markers help the readers to understand the writing from the writer's point of view.

A model of metadiscourse was suggested by [Kopple \(1985\)](#), who classified metadiscourse into two categories, namely, 'textual' and 'interpersonal'. The textual metadiscourse includes text connectives, code glosses, illocution markers, and narrators. The interpersonal metadiscourse includes validity markers, attitude markers, and commentaries. Though it was one of the first models of categorization for metadiscourse, the uncertainty of the categories evoked the need to create another model of metadiscourse.

[Hyland \(2004\)](#) further stated that interactive metadiscourse helps to organize the discourse and guides the readers throughout the text by informing the assessment of the writer. It helps the readers to understand clearly the writer's notion and interpretation. The interactive resources of metadiscourse are divided into five sub-categories, namely, transition markers, frame markers, endophoric markers, evidentials, and code glosses. Transition markers refer to a cluster of devices that includes conjunctions and conjunctives. These markers express the semantic relation between main clauses. For instance, 'in addition', 'but', 'thus', 'and', 'therefore', 'moreover', 'as a result of that', 'for example', 'however', 'nevertheless' etc. Frame markers are used to refer to the text boundaries that include items used to sequence or label text stages. It mainly helps to decorate the text and also helps to understand the topic shifts. It includes 'finally', 'to conclude', 'my purpose is to', 'first of all', 'next', 'secondly', 'lastly', 'in conclusion' etc. Endophoric markers are used to inform the writer's intention to the readers by mentioning the other parts of the text. It is used for a better understanding of the text. For example, 'noted above', 'see section', 'see figure', 'in this section', 'in this chapter' etc. Evidentials refer to the source of information from the other texts. It is used for determining authorial command by drawing information from another source. For example, 'according to X', 'X states that', '(Y, 1990)' etc. Code glosses provide the signal to abstract information. It gives additional information to the readers for

ease of understanding of the writer's notion. For example, 'namely', 'e.g.', 'such as', 'in other words' (Hyland, 2004, p. 139).

According to (Hyland, 2004, p. 139), interactional metadiscourse "concerns the writer's efforts to control the level of personality in a text and establish a suitable relationship to his or her data, arguments, and audience, marking the degree of intimacy, the expression of attitude, the communication of commitments, and the extent of reader involvement." It is divided into five sub-categories: hedges, boosters, attitude markers, engagement markers, and self-mention. (Hyland, 2004, p. 139) stated that "hedges mark the writer's reluctance to present propositional information categorically". These metadiscourse features are used to restrict the writer's full commitment to statements by showing some doubts. It includes 'might', 'perhaps', 'possible', 'may be', 'possibly', 'probably', 'likely', 'assume', 'doubt', 'apparently', 'from my point of view', 'in my opinion', 'suggests' etc. Boosters are used to highlight the writer's certainty in a proposition. These features are used to establish the writer's firm opinion. For example, 'in fact', 'definitely', 'it is clear that', 'certainly', 'definitely', 'clearly', 'beyond doubt', 'we proved', 'we found', 'substantially' etc. (Hyland, 2004, p. 139) stated that attitude markers "express the writer's appraisal of propositional information, conveying surprise obligation, agreement, importance, and so on." It refers to the features that uphold the writer's assumption about a proposition. For example, 'unfortunately', 'surprisingly', 'correctly', 'rightly', 'hopefully', 'fortunately', 'I am surprised', 'I agree', 'X claims that' etc. Engagement markers are used to build an explicit relationship with the readers. These markers are used to involve the readers in the text. Some examples of engagement markers are 'consider', 'note that', 'you can see that', 'we', 'our', personal pronouns, etc. According to (Hyland, 2004, p. 140), "self-mentions suggest the extent of author presence in terms of first-person pronouns and possessives." These features reveal the presence and participation of the author in the text in front of the readers. It includes 'I', 'the author', 'our', 'we' etc. The functions of the interactive and interactional metadiscourse markers are depicted in Table 3.1.

Table 3.1: A model of metadiscourse in academic texts by Hyland (2004, p. 139)

	Category	Function	Examples
Interactive	Transitions	Express semantic relation between main clauses	In addition/but/thus/and
	Frame markers	Refer to discourse acts, sequences, or text stages	Finally / to conclude / my purpose is to
	Endophoric markers	Refer to information in other parts of the text	Noted above / see Fig. / in Section 2
	Evidentials	Refer to source of information from other texts	According to X / (Y, 1990) / Z states
	Code glosses	Help readers grasp meanings of ideational material	Namely / e.g. / such as / in other words
Interactional	Hedges	Withhold writer's full commitment to proposition	Might/perhaps/possible/about
	Boosters	Emphasise force or writer's certainty in proposition	In fact definitely / it is clear that
	Attitude markers	Express writer's attitude proposition	Unfortunately / I to agree / surprisingly
	Engagement markers	Explicitly refer to or build relationship with reader	Consider / note that / you can see that
	Self-mentions	Explicit reference to author(s)	I / we / my /our

Hyland (2005a, p. 43) presented another model of metadiscourse category where the term metadiscourse is described as “the cover term for the self-reflective expressions used to negotiate interactional meanings in a text, assisting the writer (or speaker) to express a viewpoint and engage with readers as members of a particular community”. In this study, the metadiscourse model is divided into two major categories: textual (interactive) metadiscourse and interpersonal (interactional) metadiscourse. Whereas the interactive metadiscourse category deals with the writer’s endeavor to draw the interest of the readers by presenting a logical and rational view of the argument, the interactional metadiscourse category attempts to uphold the writer’s endeavor to give the readers a clear vision of his thoughts and also involving them in the text.

Metadiscourse is not only confined to verbal language but also expanded its limit beyond that. Apart from the verbal and written language, human beings communicate in various ways too. Human communication expands its area from verbal and written language to the visual realm. Thus, the need for visual metadiscourse evokes an understanding of the visual needs of the readers. Kumpf (2000) proposed the concept of visual metadiscourse that describes while writing, the authors have to remember the visual needs of the readers. According to Kumpf (2000, p. 403), “a need for metadiscourse in the visual realm in technical writing has grown due to the influence of computers and their increased graphics capabilities”. The author also puts forward that although during the age of typewriters, the capability of the writers to use visual ornamentation was limited, with the emergence of computers, the scenario was changed. Kumpf (2000) also discussed the categories of metadiscourse that follow: first impression, heft, convention, chunking, external skeleton, consistency, expense, attraction, interpretation, and style.

In this dissertation, as our main focus is on the metadiscourse markers (interactive and interactional) used by the readers in the digital comments, therefore we restrict our study within the scope of readers’ comments and provide only a brief description of visual metadiscourse.

3.3 Recent Studies

A considerable amount of research studies has been carried out in the field of metadiscourse, including interactive and interactional metadiscourse. In the study of [Abdi \(2002\)](#), two major academic fields, social sciences, and natural sciences, were selected. From both fields, a total of 55 academic research articles were selected. This study showed the use of interpersonal metadiscourse markers by the writers of academic research articles from social science and natural science. It mainly studied the use of hedges, emphatics, and attitude markers. The analysis showed that social science writers use the interpersonal metadiscourse markers more frequently in comparison to natural science writers. Although it further displayed that a significant difference was noticed in the use of hedges and attitude markers by the social science and natural science writers, very little difference was noticed in the use of ‘emphatics’ by the writers of both fields.

[Hyland and Tse \(2004\)](#) presented a study where the authors analyzed 240 postgraduate dissertations of L2 students, which includes 4 million words. The corpus consisted of a total of 40 dissertations that, include 20 dissertations from the Ph.D. students and 20 dissertations from the master’s students. It also included six academic disciplines, namely, Electronic Engineering (EE), Computer Science (CS), Business Studies (BS), Biology (Bio), Applied Linguistics (AL), and Public Administration (PA). The authors intended to present a robust metadiscourse model and explain how this model can help the writers for a better understanding of metadiscourse. The authors followed the metadiscourse model of [Hyland \(2004\)](#) to analyze the text. The result of the analysis showed that the post-graduate writers used interactive metadiscourse markers slightly more in their writings in comparison to the interactional metadiscourse markers. Apart from that, hedges and transitions were most frequently used by the writers in their writings which is followed by engagement markers and evidentials. This study pointed to the fact that there are variations in the use of metadiscourse devices among Ph.D. and master’s students. The main reason for this variation is the length of the Ph.D.

corpus, which was twice as long as the master's corpus, and this encourages the Ph.D. students to use more interactive metadiscourse devices in their writings to structure the elaborated arguments. This study revealed that metadiscourse is the link between the text and disciplinary cultures. However, the difference in the use of metadiscourse devices by the writers indicates an important medium to differentiate various discourse communities.

Blagojevic (2004) studied the use of metadiscourse where the author selected academic articles written in English by English and Norwegian writers from three areas, namely, Sociology, Psychology, and Philosophy. The results showed a considerable difference among the writers of the aforementioned disciplines in the use of metadiscourse. It was reported that attitude markers were used widely by the writers of Psychology, whereas the writers of philosophy were concerned with direct statements. Regarding diversity, the writers of philosophy were in the leading position in comparison to the writers of other disciplines.

Rahimpour and Faghih (2009) presented an analysis of a corpus consisting of ninety discussion sections from different research articles on applied linguistics. The corpus consisted of three types of texts, namely, English texts written by native speakers of English, English texts written by Iranian writers who are non-native speakers of English, and Persian texts written by Iranian writers. The author focused on examining the interactive and interactional metadiscourse markers. The interactive metadiscourse markers include transitions, frame markers, endophoric markers, evidentials, and code glosses. The interactional metadiscourse markers include hedges, boosters, attitude markers, engagement markers, and self-mentions. The authors adapted the metadiscourse model suggested by Hyland (2004) for this study. After investigating the metadiscourse types, a chi-square test was carried out, and it was seen that the influence of the mother tongue was responsible for the difference in the use of metadiscourse among these writers. A significant difference was noticed in the selection of the metadiscourse markers by both groups of writers, where the use of interactive metadis-

course markers was predominant in comparison to the use of interactional metadiscourse markers.

As most of the concentration of the study of metadiscourse was on academic writing, [Ädel \(2010\)](#) in her study focused on spoken and written types of metadiscourse. The principle aim of this study was to present an empirically based taxonomy of spoken and written metadiscourse. It compared the use of personal metadiscourse in 30 university lectures and 130 essays of graduate students. The lectures and essays were based on academic English. For investigating the distribution of the discourse devices in both the spoken and written mode, one taxonomy was proposed that features both the similarities as well as the differences. The proposed taxonomy contained 23 discourse functions that were divided into four main categories, namely, metalinguistic comments, discourse organization, speech act labels, and references to the audience. The outcome of this study suggested that in both the spoken and written forms, the discourse functions occurred. However, a greater range of discourse actions was present in the spoken metadiscourse rather than in the written metadiscourse. This study explored some factors that are responsible for the difference in the use of metadiscourse in both the spoken and written mode. For example, in the spoken mode, the lack of time for planning and rehearsing the speech and the presence of the audience were responsible, whereas, in the written mode, the use of metadiscourse devices was mostly common while arguing. In addition, factors related to the genre are also responsible for the variation in the use of metadiscourse devices in the spoken and written form.

Earlier, the term metadiscourse was thought of as a way of understanding language and its use, gradually this term proves to encompass a wide area in which the writers organize their texts, include themselves as well as the readers in the text, and convey their own attitudes towards the text and audience. [Hyland \(2010\)](#) examined a corpus consisting of 240 dissertations of L2 doctoral and master's students. The corpus contains a total number of 4 million words. Six academic disciplines were used to evaluate the use of resources by the L2 writers that, include Electronic Engineering (EE), Com-

puter Science (CS), Business Studies (BS), Biology (Bio), Applied Linguistics (AL), and Public Administration (PA). For the corpus, 20 doctoral and 20 master's dissertations were collected from the students who were from five Hong Kong universities, and they use Cantonese as their first language. Among the four million words, 2.6 million words were from the Ph.D. students, and 1.4 million words were collected from the texts of the master's students. The author used a text analysis and concordance program called MonoConc Pro for analyzing the corpus. The study revealed that the interactive metadiscourse markers were used more than the interactional ones. It had been noticed that among the metadiscourse markers, hedges and transition markers were most frequently used by the writers. The result of this study is presented in Table 3.2.

Table 3.2: Study of Hyland (2010, p. 132)

	Category	Masters	Doctoral	All
Interactive	Transitions	75.8	95.6	89
	Evidentials	40	76.2	64.1
	Code glosses	27.4	40.6	36.2
	Frame markers	20.7	30.3	27.1
	Endophorics	22.3	24	23.4
	Total	186.1	266.7	239.8
Interactional	Hedges	86.1	95.6	92.4
	Engagement markers	39.7	51.9	47.8
	Boosters	31.7	35.3	34.1
	Attitude markers	20.4	18.5	19.2
	Self mentions	14.2	40.2	31.5
	Total	192.2	241.5	225

A study by Pérez-Llantada (2010) analyzed metadiscourse in two important research article sections: introductions and discussions. For the corpus, 144 introduction and 144 discussion sections were selected from the Spanish-English Research Article Corpus. All the introduction and discussion sections were based on the biomedical component, and the corpus was divided into three sub-corpora: the ENG sub-corpora, the SPENG sub-corpora, and the SP sub-corpora. Whereas the ENG sub-corpus comprised of 48 research articles written in English by North American-based scholars, the

SPENG sub-corpora consisted of 48 research articles written in English by Spanish scholars. The SP sub-corpora contained 48 research articles written by Spanish scholars. The articles of ENG and SPENG sub-corpora were extracted from a sample of high-impact factor international journals. The articles of SP sub-corpora were collected from a sample of Spanish journals. The result of the study showed that the average frequencies of the metadiscourse markers are almost similar in the two article sections. The use of metadiscourse revealed various cultural and language-specific features. It also revealed the inclination for impersonal metadiscourse types and various textual development for composing arguments.

In the same year, another study was conducted by Bondi (2010) that analyzed a corpus that consisted of article introductions and a corpus of textbooks that contained introductory chapters. Both the corpora were extracted from the same domain, namely, economics. In both the introduction sections of the article and the textbook, framework sequences were studied that further pointed out discourse through a combination of verbal and nominal elements. The study aimed to examine how academic arguments in different genres are represented by using some English metadiscursive expressions. These metadiscursive expressions include forms of self-mentions and elocution markers. The study intended to uphold the cross-generic variation and explain the portrayal of disciplinary discourse in research and didactic genre. It reported both the similarities and differences between the genres regarding the collocational and phrasal patterns.

Gillaerts and Van de Velde (2010) in their work emphasized some research articles from the domain of applied linguistics to analyze interactional metadiscourse markers. Following the metadiscourse model of Hyland (2005a), this work focused on the interactional metadiscourse markers, namely, boosters, hedges, and attitude markers, in a survey corpus of abstracts of the research articles. The result of the study showed the gradual changes that took place in the use of interactional metadiscourse markers in the abstracts of the research articles.

Mur-Dueñas (2011) provided a slightly different model of metadiscourse in the study

where the author has categorized metadiscourse into two main categories, namely, interactive metadiscourse and interactional metadiscourse. The sub-categories of interactive metadiscourse are seen to differ from the previous models, whereas the categories of the interactional metadiscourse remain the same. The interactive metadiscourse category includes logical markers, code glosses, sequencers, topicalisers, endophoric markers, and evidentials. The author focused on the difference in using metadiscourse features that were used in different research articles on Business Management from different contexts, namely, North American and Spanish. The study works on analyzing a corpus that consists of 12 research articles published in international journals by the scholars of a North American University and 12 research articles published in the national journals by the scholars of a Spanish University. The research articles were written in English by the scholars of North American University, whereas the scholars of Spanish University wrote the research articles in Spanish. In the case of using metadiscourse features, a remarkable amount of difference was seen in the writings of the American and Spanish scholars. The cultural context of the publications implied great effect regarding the selection of the metadiscourse features in the writings of the research scholars.

In the study of [Sultan \(2011\)](#), metadiscourse expressions were analyzed. The author focused on analyzing a corpus of seventy discussion sections taken from linguistics research articles from English and Arabic international academic journals. These articles were written by native speakers of English and Arabic. This study focused on understanding cultural differences by analyzing the interactive and interactional metadiscourse expressions. In order to understand the difference in the use of metadiscourse expressions by the two groups of writers, chi-square tests were carried out. The author followed the metadiscourse model of [Hyland \(2004\)](#) to analyze the corpus. The study suggested that the interactive metadiscourse expressions were used by both the writers of English and Arabic in comparison to the interactional metadiscourse markers. It also suggested that Arabic writers were more inclined to use metadiscourse markers in

their writings. There were different factors that were responsible for the selection of metadiscourse markers among the American and the Iranian writers, which includes culture, genre, and the experience of foreign language of the Iranian writers.

Zarei and Mansoori (2011) in their study discussed the use of metadiscourse in the two general streams, namely, applied linguistics and computer engineering. Whereas applied linguistics represents the humanities discipline, computer engineering represents the non-humanities discipline. The study is based on the Persian and English languages. For the analysis, the corpus was formed by selecting some random articles from English and Persian writers. The articles were written by native speakers of English and Persian language. The authors used the metadiscourse model suggested by Hyland and Tse (2004). The results of the analysis pointed out that the use of metadiscourse was higher in the Persian articles in comparison to the English ones. The metadiscourse elements were frequently present in the applied linguistics discipline in comparison to the computer engineering discipline. The Persian writers used interactive metadiscourse elements more instead of using interactional metadiscourse elements. The study also put forward the fact that the cultural background of the writers used to have a great influence on the writings of the English and Persian authors. It suggested that the Persian writers were inclined more to textuality compared to the English writers. Again, whereas the English writers focused on the interactional metadiscourse, the Persian authors were less concerned about the interactional metadiscourse.

In 2012, a study by Carrió-Pastor (2012) explored the use of epistemic expressions in scientific English. The study also investigates if native speakers of English use epistemic expressions in the same way as non-native speakers of English. The corpus of this study comprised 50 research papers that are written by native English speakers and 50 scientific papers written by non-native speakers (Spanish speakers) of English. This study concentrated on the frequency of the use of epistemic modal verbs in the writings by both native and non-native speakers of English. The corpus was analyzed

by using WordSmith tool 5.0. The outcome of the study showed that although the intention of the writers to use epistemic modals in their writings was the same, there was a considerable amount of difference. The difference was noticed in the use of the modals like ‘can’, ‘may’, ‘must’ etc., in the writings of both the native and non-native writers. whereas the native speakers used ‘may’ in their writings more in comparison to the non-native writers, the non-native Spanish speakers used ‘must’ and ‘can’ more in their writings.

A study carried out by [Vahid Dastjerdi et al. \(2012\)](#) attempted to examine the impact of explicit instruction of metadiscourse markers in the writings of advanced, intermediate, and elementary EFL learners. For conducting the study, 92 undergraduate students were selected as participants from the University of Isfahan. Among them, 32 were from elementary, 32 were intermediate, and 30 were advanced students who were majoring in English literature. These participants were divided into three levels according to language competence, followed by a Michigan proficiency test. In order to extract the authentic data, the participants were put in a test to verify their knowledge along with their ability to use the metadiscourse markers in the writings. Six successive sessions were carried out to make the participants understand the explicit use of metadiscourse markers in writing. After this, a post-test was organized to examine the writing ability of the participants using the metadiscourse markers. The outcome of the test showed that the students belonging to the intermediate level had improved significantly in comparison to the elementary and advanced students. The findings of the study suggested that metadiscourse awareness had a great impact on the learner’s language performance. Thus, this study asked the teachers, practitioners, and researchers to pay more attention to metadiscourse markers in making an EFL curriculum. It also advised reinforcing metadiscourse markers through explicit instruction in EFL courses for the sake of improvement in the writing of the students.

[Pooresfahani et al. \(2012\)](#) investigated engineering and applied linguistics domains to study the use of interactive and interactional metadiscoursal features. A corpus of

eight research articles from each domain was selected, and the analysis was carried out by following the model of Hyland (2005a). The outcome of the analysis showed that in both domains, interactive metadiscourse was widely used by the writers in comparison to interactional metadiscourse. A remarkable amount of difference was noticed regarding the frequency of the metadiscoursal features among the writers of both domains. The outcome of this study upheld the fact that in both domains, the writers were eager to help the readers to understand the text rather than including them in the text itself.

In the past few decades, extensive research has been carried out in interactive and interactional metadiscourse. Cao and Hu (2014) studied the quantitative and qualitative research articles from the disciplines of education, applied linguistics, and psychology. A corpus of 120 research articles was prepared to study interactive metadiscourse by following the metadiscourse model of Hyland (2005a). This work concentrated on several aspects of cross-paradigmatic and cross-disciplinary differences in the quantitative and qualitative research articles.

In a recent paper by Salek (2014), an extensive study was made on English research articles where it was stated that the use of metadiscourse markers plays an important role in enhancing the writing skills of the researchers. In this regard, the use of metadiscourse markers in various English research articles was examined, starting from the abstract to the conclusion section following the model of Hyland (2005a) and the strategies of Abdi et al. (2010). This work showed that interactional markers such as boosters, attitude markers, and hedges were widely used in the discussion and the conclusion sections of the research articles.

The study of Alshahrani (2015) presented a comparative study of the interactive metadiscourse in academic writings. The writings were selected from English native speakers and doctorate students of linguistics who were native speakers of Arabic. The author followed the model of Hyland (2005a) to find out the similarities as well as the differences between the native English writers and the native Arab students. The

corpus of this study comprises 80 discussion and conclusion chapters. For analyzing the data, the author used Wordsmith (v. 6.0.0.186, Scott, 2012) tool. For comparing the metadiscourse categories in the writings of the English and Arab students, chi-square statistical analysis tests were performed. The result of this study suggested the influence of local culture in the writings of the Arab students that led to the frequent use of interactive metadiscourse in their writings.

In 2020, GezegİN and Melike (2020) conducted a study of metadiscourse used in academic writing where the authors showed the comparison of the use of interactional metadiscourse markers in research articles and book reviews. This study was carried out to find out the similarities and differences in the linguistic preferences of the authors. The corpus of this study consisted of 24 research articles and 24 book reviews by Turkish academic writers of English based on different disciplines. All the research articles and book reviews were written from 2004 to 2016. The authors focused on the conclusion part of the two genres. The corpus was annotated by using the UAM CorpusTool, and the authors followed the taxonomy of metadiscourse proposed by Hyland (2004, 2005a). The results of this study showed that in both the sub-corpora interactional metadiscourse markers (hedges, boosters, attitude markers, self-mentions, and engagement markers) were used by the writers. However, a remarkable difference was noticed regarding the authorial stance of the writers with interactional metadiscourse markers. The outcome of the study suggested that attitude markers were enormously used by the writers of the book reviews, who were more evaluative in their conclusions. On the other hand, the writers of research articles used hedges in their writings that showed the authors' cautiousness towards their commitment to the propositions.

Liao (2020) explored the use of metadiscourse devices by L2 Chinese writers with different writing proficiency. This study investigated how these writers use metadiscourse in their descriptive writings and how the metadiscourse devices create a difference regarding the levels of proficiency among the writers. For the experiment, 62 L1 Chinese English college learners from the United States were selected who were in

China on a study-abroad program at the time of data collection. The dataset of this study consisted of 62 descriptive Chinese essays. The participants were from the 19-22 years of age group, and there were 27 females and 35 males. A descriptive writing task was conducted on both the lower-level and higher-level learners. For analyzing the data, the authors followed the metadiscourse model of Hyland (2005a). The outcome of the study suggested that L2 learners with different writing proficiency use metadiscourse devices in their writing differently. The more proficient writers used metadiscourse devices more in comparison to the low-proficiency writers especially frame markers and engagement markers. The study also hinted at the positive relationship between advanced writers with various linguistic features.

Another study by Akoto (2020) investigated the similarities and differences in the use of metadiscourse markers in the masters' theses of the students from the University of Cape Coast (UCC). The Master's degree students were from the Sociology department. This study focused on the introduction and the literature review part of the masters' theses of ten students. For analyzing the corpus, the interpersonal metadiscourse model of Hyland (2005a) was followed, and the authors annotated the metadiscourse devices manually. The findings of the study suggested that there are both similarities as well as differences in the use of metadiscourse markers in the introduction and the literature review part of the theses. In both chapters, among the interactive metadiscourse markers, the frequency of transitions and among the interactional metadiscourse markers, the frequency of hedges were mostly seen. Moreover, a considerable amount of difference has been noticed in the use of interactive and interactional metadiscourse markers in both the introduction and the literature review chapters. Also, it was seen that in the Literature review chapter, the use of interactive metadiscourse devices was a bit more in comparison to the introduction chapter.

Carrió-Pastor (2020) identified epistemic modals used in engineering, medicine, and linguistic research papers. This study also investigated the phraseological patterns related to the epistemic modals. The corpus of this study consisted of 50 academic

research journals from each of the three domains, namely engineering, linguistics, and medicine. For analyzing the corpus, a specific tool named Metool that identifies epistemic expressions was used. This tool was developed as a part of a research project funded by the Ministerio de Economía y Competitividad, Spain. The findings of this study showed that there is a considerable amount of difference in the use of epistemic modals in the academic journal of engineering, linguistics, and medicine. It also pointed out the fact that authors who belong to various particular fields of knowledge use dissimilar epistemic modals in English. It also suggested that modal auxiliaries are the epistemic modals that are frequently used by the writers of engineering, medicine, and linguistics, respectively. In comparison to medical doctors and engineers, linguists used modal adjectives and adverbs more in their writings. On the contrary, it has been noticed that among the engineers, medical doctors, and linguists, the medical doctors understand more the importance of epistemic modality, and for expressing certainty and possibility, they use certain devices in their writings.

In order to write good persuasive writing, it is necessary to understand the effective use of metadiscourse markers. According to [Korau and Aliyu \(2020\)](#), a large number of Nigerian undergraduate students used to have problems while using metadiscourse in their writings. The study of [Korau and Aliyu \(2020\)](#) investigated the use of metadiscourse in the persuasive writings of Nigerian undergraduate students. The basis of the investigation was the relationship between the frequency of the use of metadiscourse in the writings and the persuasive writing quality. The participants were second-year English students from a Nigerian university. All the participants were asked to write persuasive essays, and after that, the metadiscourse markers were highlighted that were used in the writings. The authors followed the metadiscourse model of [Hyland \(2005a\)](#). The outcome of the study suggested that the participants used the metadiscourse markers less in their writings. The metadiscourse markers that were frequently used in the writings were frame markers and transition markers. However, the metadiscourse markers that were less frequently used were hedges, followed by engagement markers and

self-mentions.

Over the last decade, the researchers also paid attention to analyzing the metadiscourse markers in newspapers and news articles. Dafouz-Milne (2008) ventured to illuminate the metadiscoursal features (textual and interpersonal markers) used in the opinion columns of the two famous newspapers: the British newspaper named “*The Times*” along with the Spanish newspaper “*El País*”. For the investigation, a corpus of 40 opinion columns, including 20 each from both the English and Spanish newspapers, had been collected. The outcome of this study suggested that both textual as well as interpersonal markers were present in the columns of the British and Spanish newspapers, although with some variations. The study also revealed the presence of internal variations in the columns of both the English and the Spanish newspapers.

A considerable amount of research has been carried out on metadiscourse by Persian and Arabic researchers who concentrated on analyzing the use of metadiscourse markers in Persian or Arabic newspapers along with English newspapers. An interesting work by Yazdani et al. (2014) focused on analyzing the English and Persian news articles that were based on the particular event of 9/11. The author selected the mostly read newspapers from the United States and Iran to conduct this study. The articles selected from the online newspaper archive of the US were written in English by American writers, whereas the articles selected from the Persian newspapers were written by native Persian-speaking writers. The work presented the role that the interactional metadiscourse markers play in the news articles about 9/11. Following the metadiscourse classification model of Hyland (2005a), a corpus of 30 news reports was prepared to consist of 15 news articles from both English and Persian newspapers. The use of interactional metadiscourse by the English and Persian writers was depicted in Table 3.3.

The result of this study hinted at the frequent tendency of using interactional metadiscourse markers in news articles, mostly among English reporters in comparison to Iranian reporters. It also suggested that whereas the attitude markers were fre-

Table 3.3: Interactional metadiscourse study by Yazdani et al. (2014, p. 432)

Markers	Frequency		Percentage	
	English	Persian	English	Persian
Hedge	58	21	38%	32%
Booster	28	13	18%	20%
Attitude markers	40	31	26%	48%
Self-mention	18	0	12%	0
Engagement markers	9	0	6%	0

quently used by the Persian reporters in the news articles, the use of hedges was seen to be more frequent in the writings of the English writers. Overall, this study established the differences in the use of interactional metadiscourse markers in the writings of English and Persian journalists.

Kuhi and Mojood (2014) in their study prepared a corpus of 60 newspaper editorials from 10 famous newspapers of America and Iran to analyze the distribution of the metadiscourse markers. Among these newspaper editorials, 30 editorials were written in English, whereas the rest of the 30 editorials were written in the Persian language. This study followed the metadiscourse model proposed by Hyland (2005a). Here the authors analyzed the interactive and interactional metadiscourse markers from English and Persian newspaper editorials. The result of the study suggested the crucial role of genre conventions that led to the similarities in the selection of metadiscourse markers by both the English and Persian writers. It presented the similarities as well as the differences in the use of metadiscourse markers by the English and Persian journalists. In both the newspaper editorials, it was seen that in comparison to the interactive metadiscourse markers, the interactional ones were more often used by the English and Persian journalists. It also suggested that the interactional metadiscourse markers and attitude markers were proved to be the most crucial category and sub-category of metadiscourse used by newspaper editorial writers. Table 3.4 depicts the metadiscourse resources used by the writers in the English and Persian editorials.

Boshrabadi et al. (2014) in their study analyzed the role of metadiscourse markers in economic news reports. For this study, the authors selected English and Persian

Table 3.4: Findings of [Kuhi and Mojood \(2014, p. 1050\)](#)

Metadiscourse Resources	English		Persian		
	Percent	Frequency	Percent	Frequency	
Interactive	Transitions	19.5	21.91	26.32	28.33
	Frame markers	0.63	0.71	3.09	3.32
	Endophoric markers	0	0	0.21	0.23
	Evidentials	6.46	7.26	4.51	4.85
	Code glosses	8.8	9.89	4.04	4.35
	Total	35.39	39.77	38.17	41.08
Interactional	Hedges	19.84	22.3	9.68	10.42
	Boosters	10.92	12.27	14.54	15.66
	Attitude markers	27.56	30.97	29.77	32.04
	Self-mentions	0.17	0.19	1.35	1.45
	Engagement markers	6.12	6.87	6.49	6.99
	Total	64.61	72.6	61.83	66.56

economic news reports from the leading newspapers of Iran and the United States. 5 news articles from each country and a total of 10 news articles were randomly selected and analyzed in this study. The authors followed the metadiscourse model of [Kopple \(1985\)](#) to examine the functions of the metadiscourse markers in economic news reports. The study focused on analyzing the similarities and differences between English and Persian economic news reports based on the frequency of textual and interpersonal metadiscourse markers. The result showed that both the Persian and American writers used textual metadiscourse markers more in their writings in comparison to the interpersonal markers. However, it was evident from this study that interpersonal markers were used more in the writings by the American writers in comparison to the Persian writers.

[Mardani \(2017\)](#) presented an analysis of the textual and interpersonal metadiscourse markers that were used in the translated and non-translated versions of persuasive text. For this study, the author focused on the Persian-translated version of the *New York Times* along with one of the Iranian newspapers named *Hamshahri*, which was written in the Persian language. At first, 120 articles were collected from the online archives of the *New York Times* and *Hamshahri*. After that, 6 texts from each of the newspapers

were finally selected for analysis. The topic of these texts was focused on health and environmental issues. The author used the metadiscourse model of Hyland (2005a) to analyze the text, and the outcome of the analysis is presented in Table 3.5.

Table 3.5: Findings of Mardani (2017, p. 76)

	Categories	Persian originals		Translations	
		Total	Percentage	Total	Percentage
Interactive	Transitions	576	46.33	411	39.03
	Frame markers	33	2.65	52	4.93
	Endophoric markers	11	0.88	20	1.89
	Evidentials	59	4.74	97	9.21
	Code glosses	51	4.1	42	3.98
Interactional	Hedges	311	25.02	237	22.5
	Boosters	70	5.63	56	5.31
	Attitude markers	63	5.06	41	3.89
	Engagement markers	17	1.36	3	0.28
	Self-mentions	52	4.18	94	8.92
	Sum	1243		1053	

The outcome of this study upheld the fact that both the textual and interpersonal metadiscourse markers were used in the text by both groups of writers. However, it was seen that the Iranian translators used interpersonal metadiscourse markers in their writings more in comparison to the Iranian translators. This study also revealed the important role of the metadiscourse markers, which eventually helped the translators to understand the original text.

A study carried out by Noorian and Biria (2017) discussed the interpersonal metadiscourse markers in persuasive writing. The authors selected two elite newspapers, namely, ‘The New York Times’ and ‘Tehran Times’ from Iran and the United States. For the analysis, six opinion articles from each newspaper were selected. This study focused on analyzing the presence of interpersonal metadiscourse markers along with the differences in the use of interpersonal markers in the writings of American and Iranian writers. It followed the taxonomy of interpersonal metadiscourse markers proposed by Milne (2003). The result of the study showed the presence of metadiscourse categories like interpersonal markers, hedges, and attitude markers in the writings of

both Iranian and American writers.

Recent work by [Biri \(2018\)](#) focused on opinion writings on online platforms. The corpus for this study was developed by concentrating on the columns and the blog posts from various online newspapers, blogs, and news sites. The author mainly selected blogs and posts that contained issues related to ideology, politics, and society. To analyze the corpus, a multi-dimensional analysis method was used to examine the statistically significant correlation patterns between the metadiscourse markers. The outcome of the study suggested that the patterns are the representation of the relationship between the reader, the writer, and the topic. It also upheld the similarities and differences between the metadiscourse markers used by the readers in the online platform.

The use of metadiscourse devices is not only centered in the sphere of academic articles, scientific journals, newspaper reports, editorials, and newspaper articles; rather, it is also prominent in other fields as well. For example, a study by [Fuertes-Olivera et al. \(2001\)](#) discussed the use of metadiscourse in advertising. This study investigated the metadiscourse devices used by copywriters for designing slogans and headlines. The study pointed out advertising as the medium through which hidden and secret communication is shared. The authors selected some examples from a typical women's magazine and employed the Jakobsonian communication model in the discourse of advertising. In order to analyze the data, the approach of [Hyland \(1998a\)](#) was implemented. The authors followed the schema of [Hyland \(1998a\)](#) to prepare a metadiscourse schema that elaborates the function and work of metadiscourse in the headlines and slogans. The result of this study suggested that both the textual and the interpersonal metadiscourse help copywriters to convey a convincing message to the audience. The authors provided the metadiscourse schema in this study for advertising English only in the headlines and slogans. The study also showed that person markers, hedges, and emphatics are used by copywriters to caution the audience about the artificial relationship they used to have with the advertisers. It also suggested that in order to create meaningful text, endophoric markers and evidentials are used.

Kaya and Sofu (2020) focused on the role of metadiscourse markers in the writing proficiency of EFL students. For conducting this study, freshman students were taken who were majoring in teaching English as a foreign language at a university in Turkey. The authors employed one group quasi-experimental design for this study. For this study, a process-genre based writing course syllabus was selected through which the metadiscourse markers are taught. Following the one-group quasi-experimental design, a pre-test was conducted where the students were instructed to write an argumentative essay. After the training, as part of the post-test, the students were asked to write another argumentative essay on a different topic. To notice the difference between the former and later writings of the students, both texts were compared by employing a quantitative analysis method. The frequency of the metadiscourse markers used by the students in their writings was also investigated. The outcome of this study suggested that notable progress was noticed among the writers after the training, especially in the matter of organizing the text.

In 2020, a study was conducted by Rad (2020) that investigated the use of metadiscourse markers in the argumentative writings of EFL Iranian learners. For this study, 20 male and 20 female students were selected from Iranian students at Islamic Azad University (Markaz branch) and Research and Science University, respectively. All the participants were the students of English and belonged to the age group of 19-35. The participants had to perform a pretest on writing to check their initial knowledge of metadiscourse. Then all the participants were given instructions on metadiscourse markers and their use. After that, they were given passages with metadiscourse markers and were asked to identify them at first and then write their function. Finally, the participants had to perform a post-test to check their achievement in terms of metadiscourse markers. The metadiscourse model of Hyland (2005a) was followed. The findings suggested that despite being students of English, there are differences in developing the writing texts among the participants as they come from different cultural backgrounds. It was also clear from the results that both the male and the

female participants used interactive metadiscourse markers in their writings more in comparison to the interactional ones. The use of transition markers was frequent in the writings of the participants as they wanted to guide the readers through the texts. However, the instruction of metadiscourse markers helped the EFL students in their writing ability.

Another study by [Jalilifar and Alipour \(2007\)](#) examined the effect of explicit instructions of metadiscourse markers in the reading comprehension skill of pre-intermediate EFL learners. For the investigation, a total of 90 students were selected, and three groups were formed. A pre-test was conducted on the randomly chosen participants of the three groups. After that, three versions of the same text were provided to the participants. The versions were original, modified, and unmodified metadiscourse free text. The result of this study revealed that although the participants receiving the original version exceeded the participants receiving the unmodified version, the result was similar to the group receiving the modified version. A post-test was also conducted on the participants, and the results confirmed that the metadiscourse markers were primarily responsible for cohesion rather than coherence. In addition, the results revealed the positive influence of form-focused instruction of metadiscourse.

In 2020, a study conducted by [Hooi et al. \(2020\)](#) focused on the responses of the readers, especially the use of metadiscourse in online business news. For this study, 30 English major students from a public university in Malaysia were selected who were enrolled in the Corpus Studies in Language Learning course. The corpus of this study was based on business news which was selected from two prominent web portals in Malaysia: 'The Star Online' and 'Focus Malaysia'. From 'The Star Online', contemporary news was taken for the sample, and the news on new policies needed for the workplace was taken from 'Focus Malaysia'. For the annotation and analysis of the corpus, the metadiscourse model of [Dafouz-Milne \(2008\)](#) was followed that consists of textual and interpersonal metadiscourse markers. The outcome of the study suggested that most of the participants favoured reading the news with textual metadiscourse markers as

it was more comprehensible. On the contrary, 16.67 percent of the participants want to read the news without textual metadiscourse. Regarding the interpersonal metadiscourse, 93.33 percent of the participants wanted to read the news with interpersonal metadiscourse, and 6.67 percent of the participants wanted to read the news without interpersonal metadiscourse. The study also revealed that apart from business news writing, metadiscourse is crucial for writing instructors in English for Specific Purposes (ESP) writing classrooms.

Another study by [Carrió-Pastor and Muñiz Calderón \(2015\)](#) compared metadiscourse features in digital business communication. This study focused on interactional metadiscourse devices, mainly on boosters. A corpus of one hundred emails was developed by two groups of non-native English speakers who worked in an export company. Fifty emails from each group consisting of Spanish and Chinese employees were selected who use English for communicating in the business environment. The boosters were manually identified and compared with the boosters identified by [Hyland \(2005b\)](#) and [Mur-Dueñas \(2011\)](#). For the analysis of the corpus, WordSmith 5.0 software was used. The findings of this study suggested the difference in the use of boosters as the speakers were from different linguistic backgrounds. In comparison to the Chinese group of writers, the Spanish writers of English used boosters more in their business emails. This study opens the path for future research on exploring the influence of the mother tongue on English speakers.

[Huang et al. \(2020\)](#) investigated the interactional metadiscourse in English travel blogs. This study explored how travel bloggers use interactional metadiscourse in their blogs to interact with their readers. For this study, a corpus was prepared that consisted of 30 English travel blogs. The metadiscourse model of [Hyland \(2005a\)](#) was followed in this study. The frequency of the interactional metadiscourse markers was counted. The study selected qualitative and quantitative analysis to investigate the communication process of the blogger writers and the readers. This study also discussed the endeavour of the blogger writers by using typical interactional metadiscourse to inspire the reader's

future travel experiences.

Quite recently, a study by [Albalat-Mascarell and Carrió-Pastor \(2019\)](#) investigated the frequency of the self-mentions used by the candidates of the two major political parties, namely Democrats and Republicans. In this study, the investigation was carried out on the electoral debates by the two parties during the debates held for the United States presidential election of 2016. This study also analyzed how ethos or self is represented in political discourse by the electoral candidates through the use of self-mentions. The discourse was from the United States presidential election of 2016, where the Republican candidates were Donald Trump and Mike Pence, and the Democrat candidates were Hillary Clinton and Tim Kaine. The outcome of the study suggested that self-mentions were used by the Republican candidates more in comparison to the Democrats. A considerable amount of difference was noticed in the use of exclusive pronouns and self-citations by every politician to convey authority.

While in most of the studies of metadiscourse, a large corpus is analyzed, a study by [Noble \(2010\)](#) focused on small classroom-based corpus research. The author used a combination of computational and manual methods for analyzing the text that further helped the teachers to understand the competence level of the students regarding their understanding of academic argumentation. The metadiscourse model of [Ädel \(2003\)](#) was adopted by the author for analyzing the metadiscourse types from the writings of the students. This study focused on presenting a text-based teaching model that examines connectives, frame markers, code glosses, and self-mentions. The result of the analysis showed that the learners used connectives more frequently compared to the other metadiscourse markers. A remarkable difference was noticed in the high and low-scoring essays, where the metadiscourse markers were frequently used in the high-scoring essays in comparison to the low-scoring essays. On the contrary, the common English markers were used in the low-scoring essays by the writers. It was evident that the students' lack of confidence and knowledge was responsible for not being able to use the metadiscourse markers in the writings. However, this study is useful for

students to learn and be aware of text organization.

Wang et al. (2016) presented an analysis of the metadiscourse features in ten news reports based on North Korea's nuclear test in 2016. For the creation of the corpus, five news reports, each from two news websites, namely, BBC and China Daily, were collected. These news reports were reported within 24 hours after the authority of North Korea declared its nuclear test. The metadiscourse model of Hyland and Tse (2004) was adapted by the authors to analyze the news reports and UAM Corpus Tool 3.0. was used to annotate the corpus. The outcome of the study suggested the frequent use of interactional metadiscourse in the nuclear news reports more than the interactive ones. Whereas the interactional metadiscourse features were used 223 times in the nuclear news reports, the interactive metadiscourse was used 116 times (Cf. Table 3.6).

Table 3.6: Wang et al. (2016, p. 79): interactive metadiscourse in the ten reports

	Frequency	Percentage
Transition markers	113	68.07%
Frame markers	9	5.42%
Endophoric markers	1	0.60%
Evidence markers	40	24.10%
Code glosses	3	1.81%
Total	166	100%

Table 3.7: Wang et al. (2016, p. 80): interactional metadiscourse in the ten reports

	Frequency	percentage
Hedges	36	16.14%
Boosters	70	31.39%
Attitude markers	109	48.88%
Self mention	1	0.45%
Engagement markers	7	3.14%
Total	223	100%

Again, attitude markers and boosters were used frequently among the interactional metadiscourse, which pointed out the first reaction of the other countries about this event. The use of metadiscourse in the news reports presented clarity to the readers

for a better understanding of this important event.

In studying the use of metadiscoursal features used in the American and Thai English language newspapers, a Master's thesis by [Khopitak \(2015\)](#) was presented where the author studied the quantitative and stylistic features of metadiscourse used in opinion articles by American and Thai journalists. A corpus of 30 opinion articles from both American and Thai writers was prepared. In analyzing the text, the metadiscourse models of [Ädel \(2006, 2010\)](#) were used. Chi-square statistical tests were applied to test the difference in the frequencies of using metadiscourse devices by the writers. The outcome of the study threw light on the fact that although there was no notable difference, the differences in the genre led to the significant difference in using metadiscoursal features in the writings of both American and Thai journalists.

A study by [Carrió-Pastor and Alonso-Almeida \(2019\)](#) investigated the use of epistemic legitimizing strategies in online newspaper articles. These articles were based on various concepts that appeared before the independence referendum that occurred on 18th September (published from 13th to 17th September) 2014 in Scotland. This study used online newspaper articles that were published five days before the result's day. The corpus of this study consisted of 45 newspaper articles, and the epistemic stance strategies were analyzed, which were used by the journalists to justify their ideas related to the referendum. This study upheld the stance of the journalists with respect to the referendum. The corpus was manually analyzed at first, although a corpus annotating tool named WordSmith 5.0 was also used to verify the findings. The findings of this study showed the frequent use of epistemic stance strategies. The writers of the texts used epistemic modality strategies to show support in favor of independence or to oppose it. The use of epistemic stance devices in different categories was contrasted, and it was noticed that both the Yes supporters and No supporters of the campaign used epistemic modality in their writings but in different contexts.

While writing our doctoral dissertation, to the best of our knowledge, no study has been found so far on English news comments. However, [Moya Muñoz \(2016\)](#) studied

and compared the Spanish metadiscursive strategies used in news comments by the Spanish and Chilean commenters. In this study, the author analyzed the metadiscursive strategies of the digital comments in the two varieties of Spanish spoken in Chile and Spain and focused on two domains: sports and politics. The author collected digital comments from two digital newspapers named *El País* (Spain) and *Emol* (Chile). A total of 1200 comments were collected manually by the authors, and for analyzing the data, UAM CORPUS TOOL was used. In this study, the taxonomy of Hyland (2005a) was followed. The results showed that the commenters of *El País* (Spain) used metadiscursive strategies more compared to the commenters of *Emol* (Chile). This study included some other metadiscourse categories apart from following the taxonomy proposed by Hyland (2005a). Also, it discussed the importance of introducing other metadiscourse categories as the features of digital discourse are different from academic discourse.

3.4 Conclusion

The main purpose of the literature review of metadiscourse in this doctoral dissertation is to present an overview of the works that have been carried out on metadiscourse and its features so far. In order to conduct a systematic study in any field, a literature review is crucial to understand the development of the study in that existing field. Also, the theoretical background of metadiscourse devices helped us to have a clear picture of the research studies carried out in the field of metadiscourse and to interpret the results obtained from our corpus. In the next chapter, we discuss digital discourse and the studies carried out so far on digital discourse. Also, we briefly discuss social media and some of its applications.

Chapter 4

Digital discourse

4.1 Definition

Discourse refers to the ways in which people use language and present their thinking and understanding to prove their social existence. It can be presented in both spoken as well as written forms. It is the medium through which people express their views and opinions. In explaining the term ‘discourse’, (Lupton, 1992, p. 145) described it as “a group of ideas or patterned way of thinking which can be identified in textual and verbal communications, and can also be located in wider social structures”. It can be demonstrated in various forms, such as narration, argument, description, etc. It can also be manifested in the forms of newspapers and documentaries. Presently, discourse is also used in digital and social platforms (such as digital newspapers and social networking platforms) where the users convey their emotions in the form of comments. Digital interaction is gradually gaining popularity and attracting the attention of researchers. Discourse analysis attempts to understand social life by analyzing the language of the people. According to Gee (2004), discourse analysis aims to focus on how both the written and spoken language of people represents their social and cultural interest and outlook. For this reason, discourse analysis is widely studied by researchers to conduct research in various fields. The researchers can find out the lin-

guistic traits from the content with the help of discourse analysis. In this doctoral dissertation, discourse analysis plays an important role as it helps to bring out the different linguistic features of the digital comments from the digital newspapers of the UK, USA, and India.

The literature on digital discourse is discussed in Section 4.2. Subsections 4.2.1 and 4.2.2 briefly explain social media and instant messaging, as these are also a part of digital discourse. Subsection 4.2.3 describes comments on news and the literature on news comments. Finally, this chapter concludes in Section 4.3.

4.2 Studies on Digital Discourse

The extensive study on discourse and discourse analysis proves to be an important research topic to attract researchers. In this context, Michel Foucault, an eminent French philosopher, propounded some theories on discourse and discourse analysis (Foucault, 1972). According to him, discourse analysis emphasizes the power relationships in society that are conveyed through language and practices. In addition, it focuses on examining the social world that is expressed through language and often controlled by different sources of power. He argued that the foundation of society is languages which are affected by different sources of power. Although he praised the historical approaches as indisputable, he expressed his doubts regarding the methods of analysis and denied the historical belief behind the theory of structuralism. Regarding discourse, he opposed the thoughts that claim ‘discourse’ to be the only group of signs; rather, he focused on treating ‘discourse’ as actions that construct the objects of human conversations.

Later, a sub-genre of discourse analysis was introduced, namely critical discourse analysis, which mainly focuses on languages from the social perspective. According to Pedersen (2009), the main difference between critical discourse analysis with discourse analysis is that while discourse analysis focuses on language use, critical discourse

analysis investigates the differences between discursive and non-discursive practices. Critical discourse analysis mainly focuses on analyzing larger units rather than isolated sentences. In comparison to discourse theory, Pedersen (2009) stated that the main focus of critical discourse analysis is to differentiate between discourse and institutions and refer to them as two distinct social phenomena. It mainly focuses on the interaction between discourse and institution and also discusses the ideologies that are responsible for the production of power in society. From the view of Fairclough (2003), language and society are deeply connected with each other hence they should not be studied separately. According to Fairclough (2003), discourse analysis is the study of how texts function in relation to sociocultural practice. As language is inseparably connected with social life, it is important to include social contexts while studying a language. Moreover, critical discourse analysis not only focuses on analyzing verbal communication but also analyzes nonverbal communication like multimedia, films, gestures, the internet, etc. The digital comments made by the users uphold the voice of the people and society. Therefore, this doctoral dissertation ventures to analyze the digital comments of the readers of digital newspapers. In the last decades, researchers have shown their interest in discourse and discourse analysis. A wide range of works was carried out by the researchers by following the discourse analysis theory of Foucault. Similarly, various research works were carried out on digital discourse and digital discourse analysis.

In the last decade, a doctoral dissertation was presented by Ihlström Eriksson (2004) that emphasized the evolution of the online newspaper genre. The study analyzed different interviews, questionnaires, and websites in order to show the evolution process of the online newspaper genre and proposed a framework to comprehend the evolution genre of online newspapers. It was noticed that online newspapers have gradually turned into a particular digital genre. The theory of genre was used in this study, and the evolution process hinted at the change in the form and function of online newspapers. Online newspapers are stuffed with more content and images in comparison

to printed newspapers, and the most interesting feature of online newspapers is that it updates news throughout the day. This study also tried to modify the digital genre theory by describing the concepts of genre awareness and genre independence. It also ventured to establish the concept of ‘positioning,’ and a framework was also proposed to identify the genre characteristics in online newspapers.

Gjesdahl (2008) presented a thesis for the Master’s Degree Programme in English where the issue of language use in media was described. The study focused on identifying the differences in the use of language based on their cultural backgrounds in the online newspapers of different countries. It focused on studying the similarities in the use of language as well in online newspapers. Moreover, this study involved itself in analyzing the difference between the languages used in various online newspapers. In this work, the theory of discourse analysis and critical discourse analysis was used that examine 12 news articles from different online newspapers. The outcome of the study upheld the differences as well as the similarities in the use of language in different types of online newspapers.

Bergström (2008) in his study dealt with news content that is now easily accessible to readers with the help of the internet and media technology. The work analyzed a Swedish mail questionnaire survey that projects the attitudes and behaviors of the readers. It described the attitudes and activities of the readers of news content, specifically in the Swedish context. It presented the fact that a small number of people, in general, are interested in creating content on news websites; rather, the readers are mostly interested in sharing their views on various news articles. Although the platform of commenting on the news articles provided by the news media for the readers was highly appreciated, when it comes to reality, it was often seen that a small group of people usually participate in the conversation and express their opinions. The study clearly hinted at the distinction between the highly educated and less educated readers, and it was often viewed that the less educated people are more interested in commenting on the news articles, whereas the elite educated ones are eager to create their own

news content. While highly educated people use the online interaction platform as an informational resource, less educated people tend to consider the online interactive platform for mere amusement purposes.

Gerhards and Schäfer (2010) referred to the renowned normative theorist named Jürgen Habermas, who criticized the old mass media as it does not provide a platform for plural communication within the society. On the other hand, the internet proves to be a great platform that allows a great number of readers to express their views and opinions. This study presented a comparison of the online as well as the offline public sphere at three different levels focusing on their structural requirements, openness for participation, and their impact on society. It focused on human genome research that analyzes the internet and the print media of Germany and the USA. The outcome of the study showed that there is no such grand difference between the internet and print media as the conversation is one-sided and often dominated by popular and scientific actors.

Diakopoulos and Naaman (2011b) described in their work about the news sites that prove to be a great platform for the common people to discuss various issues. This study focused on the quality of the comments of the readers on various news articles. It discussed the requirements of the news commenters and the various journalistic approaches are also discussed that are necessary to maintain the quality of the comments. Regarding this study, the authors analyzed the comments of the readers based on different interviews and surveys from a website named SacBee.com. It further described how the quality of the comments depends on the utilization as well as the creation of the news information.

Digital discourse covers a wide range of topics, among which microblogging is regarded as one of the recent ones. In microblogging, a writer expresses his/her thoughts or feelings or his/her whereabouts in the form of short messages. A study by Lee (2011) focused on the text-making practices of micro-blogging. This study talks about status updates on Facebook which are based on the writings of a group of Cantonese and

English bilinguals. A total of 744 messages were collected from a group of participants, among which some are regular writers, and some keep themselves apart from updating their status for a week or longer. Then a traditional content analysis was performed for the collected status updates. Also, the linguistic and orthographic features of the status updates were gathered. The status updates, which are based on the writers' daily experiences, are regarded by the authors as situated literary practices. This study also accounts for the status updates by a pregnant woman during and post-pregnancy. It ponders upon the hybrid nature of the new media genre alongside focusing on the social implications of updating status.

An article by [Douai and Nofal \(2012\)](#) demonstrated the interaction of the local Arab people with people around the world on online platforms and how they share their views on global politics. It was noticed that with the emerging online platform, Arab people are coming forward to interact with people around the world on various issues, mostly based on the Muslim community and Islam. Nowadays, political controversies have become global issues upon which people around the globe share their views and opinions. In this work, different political issues are mentioned, such as the Swiss minaret ban, the so-called "Ground Zero Mosque" etc. The work presented that though traditional geographic boundaries exist, they cannot make a difference in the political and cultural world of the local and the global audience. It focused on the comments of the readers on various news articles from two news websites, namely, Al Arabiya.net and Al Jazeera.net. The findings of the study claimed that the new online platform proves to be a great place where Arab people can share their views and opinions and interact on global and political issues with people around the world.

[Knight et al. \(2013\)](#) discussed formality used in e-language. It focused on the difference in the levels of formality used in various modes of e-language. This study concentrated its attention mainly on the use of hedging in the different modes of e-language. Initially, a study was conducted by [Knight et al. \(2012\)](#) that investigated the frequency of pronouns and deictic markers in e-language based on the written and

spoken excerpts from BNC (The British National Corpus), which is a 100-million word corpus of written and spoken discourse in English. In the present study by Knight et al. (2013), an extensive study of Knight et al. (2012) was presented with diverting focus on the use of hedges in various modes of e-language. For Knight et al. (2013), the authors took the one-million-word CANELC corpus (Cambridge and Nottingham e-language Corpus) of digital English. The corpus includes data from online discussion boards, blogs, tweets, emails, and SMS messages based on a wide range of topics such as media, news, current affairs, academia, and education. It also covers various topics related to entertainment and lifestyle, for example, hobbies and leisure time, music, celebrity news, gossip, personal life, etc.

Nilsson and Carlsson (2014) studied the interviews of the politicians of Sweden and their views on new media and digital information technology. The study discussed the views of the media on the political strategies of politicians. It also focused on the political identities of the politicians. The politicians underwent various interviews, and in this way, the data was collected. During the interviews, two political identities were retrieved, namely, progressive political identity and authentic political identity. Surprisingly, both identities proved to be mere representatives of the interest of the common mass.

Weber (2014) stated that communication turns out to be interactive when people participate in it by sharing their valuable views. In this regard, online newspapers can be regarded as the most common platform for users to share their opinions in the form of comments. Weber (2014) focused on the theory of newsworthiness of Galtung and Ruge and analyzed that in the user comment section, the news factors of the different news articles are responsible for the participation level and interactivity of the readers. This study upheld the fact that the readers express their views on particular news, and mostly, controversial news stories grab the attention of the readers. However, the study argued that the news factors of the news articles play an important role in the interactivity and participation level of the audience.

Jones et al. (2015) discussed digital discourse and digital technologies in their work that attracts people's attention and involves the audience. This work was a compilation of 14 studies that were presented at "The Fifth International Roundtable on Discourse Analysis: Discourse and Digital Practices" in Hong Kong. It focused on analyzing the role that discourse analysis plays in human lives to comprehend digital practices. It also aimed to present the problems of the researchers in developing new theories with the help of traditional discourse analytic tools.

Nielsen (2015) in his book discussed how gradually the demand for digital media is rising and how people are gradually losing their faith in the local newspapers. According to Nielsen (2015), local newspapers are often biased as they present such news that supports a particular group of people. But in digital media, there lies a scope for people to interact with each other and share their opinions about various issues by posing a comment. The work discussed the scope of flourishing local journalism using digital media platforms.

A doctoral dissertation was presented by Suau Martínez et al. (2015), where the role of citizens in online media participation was discussed. The thesis mainly focused on the online participation of the citizens in London and Barcelona. The study aimed to examine the attitudes and behaviors of the citizens while participating in online activities, along with their behaviors and motivations in offline activities. The interactive process of the citizens in online participation is compared and categorized in this study. In this study, discourse analysis techniques were used, and it hinted that there is no such relation between active online and offline participation of the citizens, respectively. This study presented that the participants are more interested in taking part in the news media content rather than recreating the actual content and replacing the professional journalists. It also focused on the endeavor of the participants in challenging the traditional hegemony of professional journalism.

A study by Pinto-Coelho et al. (2019) analyzed the relationship between the authority that was given by the news media to the citizens and the authority that was

taken up by the readers when they expressed their views in the form of comments. This study focused on analyzing the news articles of the Portuguese National Plan for Dams of High Hydro-electric Potential (NPDHHP), along with the comments of the readers who commented on that particular news. It argued that the news coverage is eager to uphold the voices of the political actors, and it also focused on the interactional order and the stylistic features of the language used by the commenters while commenting on this specific news. The findings of this study presented how the readers and the commenters use citizen power and take part in political decision-making. It also discussed the effects of citizen power and their role in participating in the decision-making process.

A chapter by [Thurlow \(2018\)](#) was presented, which introduced computer-mediated discourse analysis or digital discourse studies. This study discussed three prime principles of digital discourse, namely, discourse, multimodality, and ideology. It showed how language is used as a metadiscursive and a metro lingual resource as well as a transmodalizing and a technologizing resource in social media.

The aforementioned studies are meant to provide an outline of the previous works on digital discourse. The studies are based on news media, news articles, and news comments of the readers on various online platforms. For this doctoral dissertation, these studies are crucial as they help to form ideas for analyzing the digital comments of the commenters from different digital newspapers.

4.2.1 Social Media

Digital discourse comprises various categories, among which social media is important to mention. The term social media encompasses a wide range of internet-based services that offers users a platform for social networking and content sharing. Nowadays, social media has become an integral part of everyday life, where people share their thoughts and feelings. In this section, we will present a brief discussion about the three most

popular social media applications; Twitter¹, Facebook², and Instagram³.

Twitter

Twitter is a renowned microblogging and social media platform where people share their thoughts in the form of messages and posts. Presently, it has millions of users from all over the world. Twitter was invented by Jack Dorsey in 2006. On Twitter, the messages sent by the users are known as tweets. There are options for the users to like, post and retweet the messages. Users who are not registered with Twitter could also read those messages that are publicly accessible. The users can tweet on any topic within a certain character limit. The users can follow each other and can also be followed. Although there are no restrictions for the users to follow someone back and using this application does not need to create any relationship among the users. If a user is following someone, he/she will get updates of the tweets of that particular person that he/she is following (Kwak et al., 2010). Twitter posts ‘trending topics’ that contain the most frequently used phrases and words that the users use in their posts. A protocol of using hashtag ‘#’ is common among Twitter users, and the trending topics are addressed with ‘#’ as well (Kwak et al., 2010).

Facebook

Facebook is the most popular social networking site that has gained more than 2 billion active users from all over the world. It was founded by Mark Zuckerberg in 2004. The users can access it from various devices such as personal computers, laptops, and mobile phones, provided those devices must have an internet connection. The users can create accounts by sharing some basic personal information such as name, date of birth, email address, gender, etc., and to access this the users need to put in a password. After accessing, the users can reach the home and profile pages. On the profile page, the

¹<https://twitter.com/>

²<https://facebook.com/>

³<https://instagram.com/>

users reveal themselves, while on the home page, the users get status updates and group joining requests from their friends (Caers et al., 2013). After the creation of the profile, users can send requests to friends. The acceptance of those requests from friends opens the option to see each others' posts and statuses. Presently, Facebook is widely used by people irrespective of age, starting from 10 years old to 80 years. In the last decade, it has become enormously popular among its users due to its unique characteristics. These characteristics include:

- A user can add a profile picture alongside a large cover photo just behind the profile picture.
- A user can share his/her thoughts and feelings in the form of posts by updating status, and the user's friends can even react or comment on those posts.
- While reacting to the posts, the users can use various emoticons that express various human emotions, for example, happy, sad, love, care, anger, surprise, etc.
- A user can reach the posts and status of his/her friends in the form of notification.
- Facebook also provides a platform to its users where they can chat with their friends personally as well as in groups. The users can reach each other through Facebook Messenger. Even the users can connect with each other via audio and video calls.
- Also, while video calling, the users can select various sticker options that could change the background as well as facial features of the users.
- There is also an option of being live on Facebook where a user can live stream his/her actions, which can be viewed by the user's friends.
- Also, users can create and join different groups where all the people share the same interests and share their thoughts.

Instagram

Instagram is a famous photo and video-sharing website which is used by more than one billion users from the entire globe. It is the brainchild of Kevin Systrom and Mike Krieger. The upsurge in popularity was noticed when famous personalities from all around the world actively joined Instagram. The users can share pictures concerning everything around the sun and comment on the photos and videos as well. These photos can be edited by using a variety of filters. On Instagram, users can follow their friends along with various famous celebrities and personalities at the same time. Also, they can share the posts publicly or could share them with previously approved users. Additionally, the option of live streaming is also present here.

As the prime concern of this dissertation is not social media and its applications, we restricted our study to Twitter, Facebook, and Instagram only.

4.2.2 Instant Messaging

As smartphones are easily accessible nowadays, instant messaging applications are gaining enormous popularity among users. These instant messaging applications allow users to communicate with friends and families, even in groups, through text, messages, voice messages, shared files, audio, videos, phone calls, etc. The main reason for the enormous recognition of these instant messaging applications is that they are easily accessible and they contain a variety of features. Examples of these instant messaging applications are WhatsApp, Telegram, and Viber.

WhatsApp

Presently, the most popular instant messaging application is WhatsApp which has more than a billion users and is owned by Facebook. Though in the initial year, WhatsApp charged a small sum of money for a subscription, from 18 January 2016, the subscription charge was completely canceled [Sutikno et al. \(2016\)](#). It provides its users a platform where they can share text messages, audio messages, various documents, free voice, and

video calls, and also exchange photos and videos. The most interesting thing is while installing this application, users could access the contacts remaining in their address book, provided the person is also using the same application on their mobile phones. The characteristics of WhatsApp Messenger are mentioned below:

- WhatsApp is accessible through its web-based application that is convenient to Windows desktop users as well as the users of Chrome web-based browser.
- It provides ads-free service to its users, and due to the enormous popularity of Facebook, WhatsApp also gained popularity very quickly among users compared to the other instant messaging applications.
- The users can communicate in groups and can send unlimited photos, videos, and text messages.
- There are also various emoticons that can be used by users while sending text messages.
- The messages that are sent by the users while going offline are saved automatically, and it is shown when the internet service is resumed (Gon and Rawekar, 2017).
- As WhatsApp uses the phone numbers that are present in the address book, the users do not have to remember any password or username.
- Not only that, but WhatsApp has also been quite helpful in an educational setting that includes language learning as well.

Discussing the positive results of using WhatsApp in language learning, Andujar (2016, p. 63) stated that “WhatsApp constitutes a powerful educational tool to encourage second language interaction among participants and its tremendous potential to activate students’ involvement remains one of the least exploited functionalities of mobile phones”. It also helps to enhance the reading ability, listening ability as well as

writing abilities of language learners. Examples of such studies are [Hazaea and Alzubi \(2016\)](#), [Fauzi and Angkasawati \(2019\)](#), [Andujar \(2016\)](#), which address the positive effects of WhatsApp in language learning, reading, writing, and speaking.

Telegram

Apart from WhatsApp, another instant messaging application that has gained popularity among its users is Telegram. It was started by Pavel Durov, a Russian-born entrepreneur, in August 2013. After its initialization, it was in major competition with the other ruling instant messaging applications such as WhatsApp, Viber, etc. [Sutikno et al. \(2016\)](#) stated that Telegram was the extensively downloaded instant messaging application in Google Play Store. It gained popularity in most of countries, including the United States and Germany ([Sutikno et al., 2016](#)). The users can access this application by their phone numbers as initial identification. After the installation of this application, the users can access the address book and those users who use the same application. Here the features of Telegram are described below:

- Like WhatsApp, Telegram also provides subscription-free service and an ads-free environment to its users.
- In Telegram, the users can create their own unique user IDs through which they can access other users directly without even accessing the user's phone number. Thus, after adding the respective user to the contact list, the phone number of that respective user will not show, which is undoubtedly a unique strategy to maintain the privacy of the users. This feature is commendable as, nowadays, privacy has become the most important concern for users.
- Like WhatsApp, users can share text messages, audio and video messages, video, and audio calls, etc., and the users also can share documents.
- The users can use emoticons and stickers while chatting.

- It can be used on multiple applications such as Windows Operating System, Android, iOS, MAC, and Windows Phone (Sutikno et al., 2016).
- Also, the communication between the users is stored in the cloud so that later the users can retrieve the communication if needed.
- In Telegram, the communication between the users can also be in groups and channels. Users who share the same interest can create a group for better communication. The messages could be seen by all the members of the respective group. For joining such groups, a unique join link is provided by the administrators of the group or a user can be invited by an existing group member.
- Likewise, in channels, users can access the messages which are sent by the administrators as the administrators only have the power to publish messages in the channel that are accessible to an unlimited number of users. However, there is no limit for the users for subscribing to the channels, and the users can reach these channels through their particular usernames or join links (Dargahi Nobari et al., 2017).
- Unlike other social media applications such as Facebook, Twitter, etc., the users of Telegram have the option to not include themselves in any personal relationship like friendship with all the users, and also the users can have their own private contact list for creating their own friend circle according to their choice (Dargahi Nobari et al., 2017).
- Here, a single user can create an unlimited number of groups and channels and can include an unlimited number of members.

Viber

Another popular instant messaging application is Viber. The owner of Viber is a Japanese multinational company named Rakuten. It has some common features similar

to WhatsApp and Telegram, for instance, the voice calling and video calling features. Even the users can create a group for chatting. The features of Viber are shown below:

- It provides public chat forums that allow chatting openly with large communities or groups where people share the same interests.
- It has a unique feature where a user can reach another user's phone number even if he/she is not using this application, though it costs additional charges.
- Alike WhatsApp and Telegram, Viber can be used on laptops and personal computers.
- The users do not need to pay any subscription charges for using this application.
- There are two modes of voice quality, the normal mode and the HD mode, for voice and video call services.
- Aside from sending photos, the users can draw with their fingers [Sutikno et al. \(2016\)](#).

Although there are other instant messaging applications available, here we mentioned the most liked ones as this is not the main concern of this research.

4.2.3 Comments on News

In this section, we discuss comments and mention the studies about comments on news.

Definition

A comment is a medium through which readers express their opinions and feelings. In digital media, comments are made in written form. In this doctoral dissertation, we discuss news comments where the readers express their emotions in the form of comments and post them to respective digital newspapers. Due to the rapid increase of digital and social media, nowadays, people are excessively interested in sharing their

views and opinions in the form of comments. The commenters of these digital and social media are irrespective of gender, age, economic status, and educational background. According to Kalogeropoulos et al. (2017), there are various types of commenters who participate frequently on digital media:

a) People who are highly active on various social media platforms also frequently express their opinions on digital news sites.

b) The followers of left-wing and right-wing parties often comment on left-wing and right-wing newspapers, respectively.

c) People who are immensely interested in hard news are often likely to share and comment on digital news and share those via email (Kalogeropoulos et al., 2017).

In this study, we have observed two types of comments:

1. *Comments made in response to the news article:* the comments made by the readers on particular news to express their views regarding that topic.
2. *Comments made in response to a pre-comment:* the comments that are made by readers in response to the comments of a reader who commented previously. The exchange of views among the readers could form a conversation where every comment is regarded as a reply. Henceforth, a reply to a commenter by another commenter is also a comment.

Studies on News Comments

In this section, we discuss the previous studies which are conducted on news comments.

Tsagkias et al. (2009) studied the volume of online news comments. The authors prepared five feature sets and reported in their study how these features performed individually and jointly on two classification tasks. The two classification tasks include- a) a binary classification which is capable of recognizing articles that are likely to receive comments b) a second binary classification which uses the output of the former classification to identify the articles having ‘low’ and ‘high’ comment volume. The

outcome of this study shows that the first binary classification performs far better in comparison to the second one.

[Diakopoulos and Naaman \(2011a\)](#) reported the connection between topicality, temporality, sentiment, and quality in news comments. For this study, a total number of 54,540 comments were collected from a well-read Californian newspaper named Sacramento Bee, SacBee.com. All the comments were written in August 2009. This study revealed the fact that both positive and negative sentiments are crucial for indicating the quality of online discourse. The study found that negativity evokes the need for moderation of the comments.

[Diakopoulos and Naaman \(2011b\)](#) discussed the quality of online news comments. According to news organizations, the quality of online news comments is of utmost importance to maintain trustworthiness within the community. This study described how the requirements of online news commenters interact with the operation of various quality management methodologies in a journalistic setting. The authors focused on newsroom interviews and surveys and characterized the comment discourse of SacBee.com. The major quality issues were identified by the authors on the SacBee.com site. The authors then discussed how the quality of discourse depends on the consumption and production of news information. This study also mentioned the motivation of the users for reading and writing news comments.

A study by [Kothari et al. \(2013\)](#) analyzed comment-tweets that are shared by Twitter users. In comment-tweets, users share their thoughts and emotions in the form of comments, which are based on some news articles. In this paper, the authors focused on the nature of comment-tweets and compared those tweets to subjective tweets. Further, a machine learning technique approach was applied to identify the comment-tweets. To evaluate the proposed approach, the authors manually created a new test set from TREC-2011 Microblog track data. The outcome of this study showed the efficiency of the proposed classification approach that successfully identified the comment-tweets with high precision.

Hille and Bakker (2014) discussed online news comments that were outsourced. This study focused on the fact that while initially, the commenters comment on specific news sites, later outsourcing of the comments to a third-party platform has become possible with the help of Facebook. It investigated the comparative effect on the quality and quantity of comments by digital media users on news sites and Facebook. For this study, an inventory of 62 Dutch media was created from regional and national newspapers that include online news sites, newsweeklies, public and commercial broadcasters, etc. For analyzing the content, the qualitative text analysis tool MAXQDA was used. The outcome of this study reveals the fact that Facebook is the primary choice by the news media, which have the tendency to outsource comments to Facebook. Also, by outsourcing the comments, the quality, as well as the quantity of the comments, are hampered.

Ziegele et al. (2014) explored the reason for some online news discussions being more interactive compared to other ones. This study argued that the discussion factors present in the news comments attract more commenters. Various qualitative interviews were organized to know the reason from the commenters, and a quantitative content analysis of 1580 user comments was conducted. The results showed that there are various factors that trigger the interactivity between the commenters. These factors are controversy, negativity, the uncertainty of discussion factors, comprehensibility, etc.

Another study was carried out on the quality of online news comments by Park et al. (2016). Nowadays, comment moderators are assigned to check the quality of the comments before publishing them online. The comment moderators discard low-quality comments (for example, aggressive comments, comments with weak grammar and sentence structure, etc.) and highlight the high-quality ones. The filtering of comments is important as it ensures the support of the news outlet's policies. In this study, a system named CommentIQ was developed, keeping in mind a user-centered design approach. This system helps the comment moderators to identify high-quality comments by employing a combination of comment analytic scores, visualizations, and

configurable UI components. The assessment of this system was made with professional comment moderators from local and national news outlets. This study also described the usefulness of this system in transforming journalistic practices through online comments.

A study by [Kolhatkar and Taboada \(2017\)](#) explored the terms — *constructiveness* (i.e., engaging, respectful, and/or informative conversations) and *toxicity* (i.e., verbal abuse, offensive and hate). In the context of filtering news comments, they investigated the relationship between constructiveness and toxicity.

Another study on the abusive language used in news comments was carried out by [Desrul and Romadhony \(2019\)](#). With the extensive use of abusive language and derogatory comments in news comments, the need for an abusive language detecting system is also in need. This study upholds an Indonesian abusive language detecting system. A total of 3184 comments were collected from various online news sites such as ‘detik’, ‘kaskus’, ‘kompas’ published between the time period of March to September 2019. The authors employed three classifiers: Naive Bayes, Support Vector Machine (SVM), and K-Nearest Neighbor for abusive language detection task and found that the SVM classifiers performed the best.

[Zannettou et al. \(2020\)](#) in their study discussed hateful discourse on news websites. For this study, 125 million comments from 412 thousand news articles were collected, which were published between July 2016 and February 2018. The articles were selected from popular news sites and are further posted on 4chan’s Politically Incorrect board (/pol/). Also, six subreddits were selected from Reddit. To understand the reason behind the hate comments on news articles, the authors used temporal analysis, user-based analysis, and linguistic analysis and analyzed the data and comments. The results showed that news related to politics and real-life events attracts more hateful comments than any other news story. Also, a considerable linguistic difference in characteristics has been noticed between the articles containing hateful comments with those that do not contain hateful comments.

News Comments on the United Kingdom, India, and the United States of America

In the literature, researchers investigated online news comments, which includes quality of comments to estimate the trustworthiness within the community (Diakopoulos and Naaman, 2011b; Park et al., 2016) or constructiveness (Kolhatkar and Taboada, 2017), sentiment analysis (Diakopoulos and Naaman, 2011a), hateful discourse (Desrul and Romadhony, 2019; Zannettou et al., 2020), the reason for attracting commenters (Ziegele et al., 2014). However, their studies did not take into consideration ethnicity profiling which could impact a lot on discourse — particularly in the writing style and use of vocabulary. To the best of our knowledge, while writing this thesis, we did not find any study on news comments which compares the discourse characteristics of news comments. Therefore, it motivated us to investigate the metadiscourse characteristics of news comments written in the same language (English) for the three countries that belong to three continents.

4.3 Conclusion

In this chapter, we illustrate digital discourse in detail. We describe the previous studies that were carried out on digital discourse. Although we did not do an in-depth study on social media, as this doctoral dissertation is more concerned with news comments, we mentioned it briefly as social media is a part of digital discourse. The literature on news comments was also depicted in this chapter, as this is the main topic of this doctoral dissertation. In the next chapter, we discuss the corpus of our study.

Chapter 5

Corpus

5.1 Introduction

The previous chapters of this study were dedicated to the research that has been carried out in the field of metadiscourse and digital discourse. The prime focus of this chapter is to present the corpus used in this doctoral study. In Section 5.2, we will describe how the data was collected and how it was preprocessed. Section 5.3 details the description of the corpus. We concluded this chapter in Section 5.4.

It is important to have a large amount of written text to determine the differences in the use of languages. The compilation of a number of written texts is called a corpus. The source of the word ‘corpus’ is actually the Latin word ‘corpus’, which means ‘body’, and the plural form of ‘corpus’ is ‘corpora’. As Dash (2008) stated in his study, the Latin word ‘corpus’ implies two different connotations in modern English: ‘corpse’ and ‘corps’. Whereas the term ‘corpse’ emerged in the thirteenth century as ‘cors’ primarily, which means ‘body’, but later by the end of the fourteenth century, it was stated as ‘dead body’. According to Crystal (2011), a corpus is a collection of linguistic information, such as written texts or transcriptions of recorded voice, that can serve as the basis for linguistic description or as a way to test linguistic hypotheses. Therefore, a corpus is a collection of written or spoken texts that are in machine-readable form and

represent the use of language. It is conducive to determining the linguistic properties as well as characteristics of a language.

According to Meyer (2002), the Brown corpus is regarded to be the first corpus in English. *W. Nelson Francis* and *Henry Kucera* played a pioneering role in creating the first English corpus in the first part of the 1960s. Previously, linguists did not accept the idea of creating a corpus for language research. As Meyer (2002) pointed out, the idea of creating the Brown Corpus was told to be ‘a useless and foolhardy enterprise’, but later it was accepted that corpus linguistics is crucial for both ‘descriptive and theoretical studies of language.’ Based on the use of languages, a corpus can be classified as monolingual, bilingual, or multilingual. A monolingual corpus includes a particular text, whereas a bilingual corpus represents the texts of two different languages. A multilingual corpus is formed by texts of more than two languages. In this doctoral dissertation, the corpus is of a monolingual type because the entire corpus is written in English, as it consists of digital comments on newspapers by commenters from the United Kingdom, India, and the United States of America.

5.2 Data Compilation

For the research problem being discussed in this study, no corpus compiled by other researchers was available. Therefore, to conduct this research, we had no choice but to develop a corpus. We manually selected the comments from each of the digital newspapers. In this study, the digital comments of the readers were analyzed, which were from three domains, namely, sports, entertainment, and two political groups: right-wing and left-wing. The study is based on the comments on news published in three newspapers published in three nations in English: the United Kingdom, India, and the United States of America. English is the crucial medium of communication in the United Kingdom and America, and it is also one of the most spoken languages in India. English is regarded as one of the official languages in India as well. The

digital comments and the replies that we collected were written in English. The reason for selecting politics, sports, and entertainment domains was that these are highly discussed topics among the common people of the UK, India, and the USA.

It is necessary to mention here that we found it challenging to gather digital comments from India in comparison to the UK and the USA, as the common people of India are not quite active on the digital platform. There are various reasons for the Indians not being active on the digital platform:

- (a) *Scarcity of internet*: The scarcity of internet in remote areas, villages, and suburbs is the prime reason that the common people are unable to use digital and social media platforms. Though it is necessary to mention here that, at present, internet access is rapidly growing in India, and it is accessing remote areas too. But it will take time for the common people to realize the importance of expressing their views in the form of comments on digital platforms.
- (b) *Age gap*: Another potential reason is the age and generation gap. The aged people mostly prefer printed newspapers to digital ones. Only the young generation of the country relies on digital newspapers, so it is likely to happen that mostly young and middle-aged people in the country comment on digital platforms.
- (c) *Language barrier*: India is a multilingual country with 23 official languages (along with the additional official language English). It is likely that in every official language, newspapers (both printed and digital) are published on a regular basis throughout the country. So, it is not always the case that people only comment on English digital newspapers.
- (d) *Preference of vernacular medium*: Another natural tendency of the common people is that they prefer the vernacular medium of language more in comparison to English as a second language. However, it is not always the case, especially among the young generation, as they tend to prefer English rather than their vernacular languages. As a result, we had to look for various news articles to

collect the comments for the Indian corpus. On the contrary, in the case of the UK and USA, where English is the first language, we got plenty of comments to create our corpus.

5.2.1 Newspaper selection

As this study is based on digital comments, it is necessary to understand the political ideology of the commenters on specific news. To check the political ideology, we selected the newspapers from two political groups: left-wing and right-wing. The corpus of this study was created from the left-wing and right-wing newspapers of The United Kingdom, India, and The United States of America.

McClosky and Chong (1985) in their journal discussed the similarities as well as the differences between the right-wing and the left-wing political parties. Both parties used to differ from each other based on their political ideas and ideologies. Whereas the left-wing political party supports liberalism, the followers of the right-wing political party are much more conservative. In most of the nations of the world, these two political ideologies exist. There are various newspapers that follow the ideologies of the left-wing as well as right-wing political parties. The newspapers that follow the ideologies of the left-wing political party are known as the left-wing newspaper, and the newspapers that follow the right-wing party are known as the right-wing newspaper. The corpora of the left-wing were collected from 7 leading left-wing newspapers of the United Kingdom, India, and America, namely, *Independent*¹, *The Guardian*², *The Hindu*³, *The Indian Express*⁴, *The Tribune*⁵, *The Huffington Post*⁶ and *The Washington Post*⁷. The corpus of the right-wing was extracted from 5 leading right-wing newspapers of the United Kingdom, India, and the United States of America that

¹<https://www.independent.co.uk>

²<https://www.theguardian.com/international>

³<https://www.thehindu.com>

⁴<https://indianexpress.com>

⁵<https://www.tribuneindia.com>

⁶<https://www.huffingtonpost.in>

⁷<https://www.washingtonpost.com>

includes *The Telegraph*⁸, *The Daily Mail*⁹, *The Times of India*¹⁰, *The American Conservative*¹¹ and *Brietbart*¹². From these left-wing and right-wing newspapers, we have selected articles related to politics, sports, and entertainment domains. We have gathered a total number of 2,034 comments along with 2,004 replies. These comments were taken from 64 news articles and from the above-mentioned 12 newspapers. The overall statistics of the left-wing and right-wing corpus are portrayed below. On the basis of popularity or being mostly read by the readers of the UK, India, and the USA, the newspapers were selected.

Table 5.1: Overall statistics of the left-wing and right-wing corpus

Comments	2,034
Reply	2,004
News	64
Newspaper	12

5.2.2 Newspaper articles selection

As explained above, for the corpus, we have selected articles from three domains: politics, sports, and entertainment. In the first step, we manually reviewed various domains from the selected newspaper list mentioned in Section 5.2.1. Later, we concluded that commenters, irrespective of every nation tend to comment more on these three domains. The comments on the articles were collected from 2016 to 2018. During that specified time period, we made a list of trending topics in our 3 selected domains. It is a well-known fact that if topics are breaking news (such as *the sudden accidental death of renowned Indian actress Sridevi*), or political big events (such as *Brexit*), the readers are highly engaged and express their opinions in the form of comments and replies in

⁸<https://www.telegraph.co.uk>

⁹<https://www.dailymail.co.uk>

¹⁰<https://timesofindia.indiatimes.com>

¹¹<https://www.theamericanconservative.com>

¹²<https://www.breitbart.com>

exchange with other commenters. The selected news headlines and their category are given below in Table 5.2 and Table 5.3.

5.2.3 Comments selection

While collecting the comments for our study, we selected a comment based on the following assumptions:

- A1: We filtered out a comment if it is too short (i.e., contains a few words), such as if it contains a single word or an emoticon or just mentions the name of another commenter.
- A2: Also, we did not take into consideration a comment if it contains code-mixing. Because our focus is solely on monolingual English text, code-mixing is out of the scope of this dissertation.
- A3: We considered a comment valid if it did not satisfy A1 and A2 and was written in English only.
- A4: We also considered a reply to a comment if it is associated with a comment that holds A3 and the reply itself contradicts A1 and A2.

While collecting the comments, we carefully checked the additional information associated with each of the comments. This additional information includes the name of the commenter and the geographical location from where the commenter made his comment. The motivation behind checking these additional information (i.e., commenter's name and place) was to identify whether the commenter was native to the place of our consideration, such as India, the UK, or the USA. Hence, we considered a comment made by a native speaker following his name and place of the comment. Therefore, as per our consideration, the commenters of the UK based newspapers were particularly from the UK. For the sake of no ambiguity, we did not consider any comment which was commented either from outside the considered country or with a surname that was not

Table 5.2: Left-wing news corpus collection

Category	Newspaper	Date	Link	Title
GB-ENT	The Guardian	10-Apr-18	https://bit.ly/3grP3Va	Netflix v Cannes: what the film festival feud means for the future of cinema.
GB-ENT	The Guardian	10-Apr-18	https://bit.ly/3pUjKXv	10 years of Lady Gaga: how she queered mainstream pop forever
GB-POLITICS	Independent	5-Apr-18	https://bit.ly/3pWM2k1	David Miliband urges Labour to back new Brexit referendum in latest intervention in UK politics.
GB-POLITICS	Independent	5-Apr-18	https://bit.ly/3wndv0J	North Korea is 'almost certain' to have ballistic missiles that could reach UK by 2019, warn MPs
GB-POLITICS	Independent	5-Apr-18	https://bit.ly/3zsYdth	Salisbury attack: Labour accuses Government of 'playing party politics' by withholding high-level intelligence from Jeremy Corbyn.
GB-SPORTS	The Guardian	8-Apr-18	https://bit.ly/2S2OuZW	Javier Hernández dents Chelsea's top-four hopes as West Ham spy safety
INDIA-ENT	The Indian Express	28-Feb-18	https://bit.ly/3gt4I85	Sridevi dies at 54, leaves India in shock
INDIA-POL	The Hindu	15-Apr-18	https://bit.ly/3gzYWjH	Kathua rape and murder: Mehbooba Mufti thanks 'national leadership'
INDIA-POL	The Hindu	14-Apr-18	https://bit.ly/2TyxZFc	Group of Supreme Court judges tried to codify duties of CJI
INDIA-POL	The Indian Express	27-Mar-18	https://bit.ly/3gomGIL	Triple talaq illegal, now SC big bench to look at polygamy, nikah halala
INDIA-POL	The Indian Express	28-Jan-18	https://bit.ly/3cKTedD	Govt to 'leave no stone unturned' for passage of Triple Talaq bill in Budget session: Ananth Kumar
INDIA-POL	The Indian Express	29-Jan-18	https://bit.ly/35qPb4	PM Modi pushes for passage of Triple Talaq Bill, calls it 'New Year's gift for Muslim women'
INDIA-POL	The Indian Express	24-Mar-18	https://bit.ly/3ciBsrC	Shoe hurled at AIMIM chief Asaduddin Owaisi in Mumbai rally
INDIA-POL	The Indian Express	11-Mar-18	https://bit.ly/3xpc4ik	Triple talaq Bill injustice to Muslim women, oppose it, says Asaduddin Owaisi
INDIA-SPORTS	The Hindu	15-Apr-18	https://bit.ly/2Tyza7A	Saina outshines Sindhu, emerges the golden girl
INDIA-SPORTS	The Indian Express	19-Mar-18	https://bit.ly/3zvvrYU	Sri Lankans find their voice, India get their twelfth man in Colombo
INDIA-SPORTS	The Indian Express	19-Mar-18	https://bit.ly/3xt3Vtr	Sachin Tendulkar's No. 10 jersey to be retired by BCCI
INDIA-SPORTS	The Indian Express	19-Mar-18	https://bit.ly/3vqHe7L	India win U-19 World Cup: From Sachin Tendulkar to Virat Kohli, who said what on Twitter
INDIA-SPORTS	The Tribune	15-Apr-18	https://bit.ly/3xiRlww	Saina beats Sindhu to clinch gold as India finish CWG campaign with 66 medals
INDIA-SPORTS	The Tribune	14-Apr-18	https://bit.ly/3vmOCKl	Haryana's Neeraj Chopra wins javelin gold at Commonwealth Games
INDIA-SPORTS	The Tribune	20-Jun-17	https://bit.ly/3gEm4h5	India-Pak rivalry: Is the thrill gone?
INDIA-SPORTS	The Tribune	14-Apr-18	https://bit.ly/2TzqAFw	Boxers Mary Kom, Gaurav Solanki claim gold in Commonwealth Games
INDIA-SPORTS	The Tribune	17-Apr-18	https://bit.ly/3xkdNWj	Jobless gold medallist Rajput looking for Haryana govt support
INDIA-SPORTS	The Tribune	16-Apr-18	https://bit.ly/3ws33VN	Gayle is back and it is bad NEWS for other teams, says Rahul
INDIA-SPORTS	The Tribune	16-Apr-18	https://bit.ly/3iJOLeY	Will work on technique to improve personal best this season: Chopra
INDIA-SPORTS	The Indian Express	18-Apr-18	https://bit.ly/3zrqsIT	Mary Kom has no intention to retire anytime soon
INDIA-SPORTS	The Tribune	18-Apr-18	https://bit.ly/3wsodCZ	Race and sport in Aussie Sin City
INDIA-SPORTS	The Tribune	18-Apr-18	https://bit.ly/3grFFXs	Axing shooting from next CWG will harm young shooters: Jitu
INDIA-SPORTS	The Indian Express	6-Apr-18	https://bit.ly/35uWvtr	Won't play in IPL even if they invite me, says Shahid Afridi
INDIA-SPORTS	The Tribune	1-May-18	https://bit.ly/3xsLTrh	Hockey India at it, again!
INDIA-SPORTS	The Tribune	1-May-18	https://bit.ly/35lmt2C	Life comes full circle after grappling with infamy
USA-ENT	The Huffington Post	30-Mar-18	https://bit.ly/3zRwDKa	Tucker Carlson Invited Fabio On His Show, And Things Got Very Weird
USA-ENT	The Huffington Post	29-Mar-18	https://bit.ly/3ihCPQB	George W. Bush's Ethics Chief Warns Trump Insiders: People Will Go 'To The Slammer'
USA-ENT	The Huffington Post	31-Mar-18	https://bit.ly/2XWDFv5	Laura Ingraham Announces Week Off Air As Advertisers Flee
USA-POLITICS	The Huffington Post	30-Mar-18	https://bit.ly/3ijeZUJ	Donald Trump Gets Mocked After Showing He Really Doesn't Know What The Post Office Does
USA-POLITICS	The Huffington Post	31-Mar-18	https://bit.ly/3AWPmAo	Jimmy Carter Zings Donald Trump: Apparently, America Wants A Jerk For President
USA-POLITICS	The Huffington Post	29-Mar-18	https://bit.ly/3kQz06F	China Warns U.S. Not To Open Pandora's Box, Unleash Trade Ills On World
USA-SPORTS	The Washington Post	25-Feb-18	https://wapo.st/3ofYzQG	Germany comes oh-so-close to its own 'Miracle on Ice,' but Russians win Olympic gold
USA-SPORTS	The Washington Post	25-Feb-18	https://wapo.st/3upPMwS	Team USA's women's hockey gold medal dazzles fans and changes the sport
USA-SPORTS	The Washington Post	4-Apr-18	https://wapo.st/3zRyE3Z	'I'll be damned if I'll allow someone to use our platform': LeBron James warns Nick Saban
USA-SPORTS	The Washington Post	3-Apr-18	https://wapo.st/3ASaWWK	James Harden dances all over the Wizards as Rockets cruise to an easy win
USA-SPORTS	The Washington Post	7-Apr-18	https://wapo.st/2Y86Z20	Anthony Rendon ejected without a word; Dave Martinez tossed for plenty of them

Table 5.3: Right-wing news corpus collection

Category	Newspaper	Date	Link	Title
GB-ENT	The Daily Mail	7-Apr-18	https://bit.ly/2ZFXXtA	TOWIE's Amber Turner flaunts her incredible figure in a skimpy red bikini as she hits the beach in Dubai
GB-POLITICS	The Daily Mail	6-Apr-18	https://bit.ly/3F3mJEc	A 'clean Brexit' could be the worst catastrophe to hit Britain for years
GB-POLITICS	The Telegraph	1-Apr-18	https://bit.ly/3F1p7eB	Revealed: Election watchdog officials face calls to resign over Brexit 'bias'
GB-POLITICS	The Daily Mail	22-Apr-18	https://bit.ly/3zUFd5Q	Former Russian spy codenamed 'Gordon' is 'identified by police as suspect in poisoning of Sergei and Yulia Skripal'... but officers fear they will never be able to catch him.
GB-SPORTS	The Daily Mail	6-Apr-18	https://bit.ly/3utu6Qj	Paul Pogba was offered to Manchester City in January, reveals Pep Guardiola as he hits back at being called a 'coward and dog' by agent Mino Raiola
INDIA-ENT	The Times of India	10-Apr-18	https://bit.ly/3mpS77t	Bharti Singh: Kapil Sharma is my guru and it makes me cry to see him in such a state
INDIA-ENT	The Times of India	10-Apr-18	https://bit.ly/2WrYb6o	'Raazi' trailer: Alia Bhatt as 'Sehmat' delivers yet another jaw-dropping performance
INDIA-POL	The Times of India	13-Apr-18	https://bit.ly/3m5UA6J	Our daughters will get justice, culprits won't be spared: PM Modi on rape cases
INDIA-SPORTS	The Times of India	12-Apr-18	https://bit.ly/3ATLDDA	CWG 2018: Gold for Sushil Kumar, Rahul Aware and silver for Babita Kumari as wrestlers dominate
INDIA-SPORTS	The Times of India	11-Apr-18	https://bit.ly/3CYEdQm	IPL 2018: Just Chill, tweets Vinay Kumar after 19-run over in KKR loss
INDIA-SPORTS	The Times of India	12-Apr-18	https://bit.ly/3CUMHHZ	IPL 2018: Rajasthan Royals beat Delhi Daredevils by 10 runs (DLS)
INDIA-SPORTS	The Times of India	10-Apr-18	https://bit.ly/3F52aY6	I see a lot of me in Virat Kohli: Ravi Shastri
INDIA-SPORTS	The Times of India	13-Apr-18	https://bit.ly/3AT8cse	IPL 2018: Sunrisers Hyderabad beat Mumbai Indians by one wicket in a last-ball thriller
USA-ENT	Breitbart	6-Apr-18	https://bit.ly/3ow5cif	Hawkins: Why Celebrities Couldn't Use the YouTube HQ Shooting to Push Gun Control
USA-POLITICS	The American Conservative	3-Apr-18	https://bit.ly/39MHfud	Jordan Peterson Claims He's No Conservative
USA-POLITICS	The American Conservative	28-Mar-18	https://bit.ly/3AVn9Kq	Trump Wants to Put Drug Dealers to Death
USA-POLITICS	The American Conservative	4-Apr-18	https://bit.ly/3kQIcbq	The GOP's Laughable Call for a Balanced Budget Amendment
USA-POLITICS	The American Conservative	4-Apr-18	https://bit.ly/3of7ZMv	Donald Trump: Master of the Masculine ShriII
USA-POLITICS	The American Conservative	29-Mar-18	https://bit.ly/3iAUyCd	Five Dems in Trump Country Key To GOP Senate Win This Fall
USA-POLITICS	The American Conservative	25-Sep-17	https://bit.ly/3kQH9bf	But things are much shakier for the Republican majority in the House.
USA-SPORTS	The American Conservative	6-Jun-16	https://bit.ly/3ochYII	Trump's Football Follies
USA-SPORTS	The American Conservative	30-Mar-17	https://bit.ly/3F2pIC7	The Trans-formation Of Women's Sports
USA-SPORTS	The American Conservative	30-Mar-17	https://bit.ly/3F2pIC7	Sports: Red America's Achilles Heel

common to that country. We leveraged the name and surname list from *Wikipedia*¹³ for validating the name of the commenters under human supervision.

5.2.4 Data pre-processing

After collecting the data, the next step was to exclude the irrelevant elements from the corpus. We manually checked each collected comment and cleaned the data as per the followings:

1. Our dissertation is based on digital comments that we collected from the mostly read digital newspapers in the UK, India, and the USA. We wanted to anonymize the comments so that they would not reveal the identity of the commenters.

¹³<https://bit.ly/3NeHuDa>, <https://bit.ly/44hITiL>, <https://bit.ly/41WKisQ>

Moreover, replacing the names with some implicit names would not affect the identification of metadiscourse markers. Therefore, we have not considered the *name of the commenters* to be of prime importance; instead, we replaced the names by marking them as *@COMMENTER*.

2. We replaced the *links of the websites* by *URL*, which were used by the commenters in their comments, as the links are not useful for metadiscourse markers as well as they could project the author's identity.
3. As we considered only text, *emojis* and *emoticons* were removed from the comments of the commenters.
4. While we collected the comments, the *date and time* information was associated with each comment. As for our study, this information was not required, we deleted the date and time information from each comment.
5. Also, we observed some *garbage characters* in the comments. To facilitate the smooth processing of the text with a tagging tool, we deleted them.

5.3 Corpus Description

The corpus is divided into two sub-corpora: left-wing and right-wing, and it focuses on mainly three domains: politics, sports, and entertainment. The digital comments were collected from the left-wing and the right-wing newspapers of the three nations. Although among the newspapers, we thought to select the internationally acclaimed ones, due to the lack of digital comments from the internationally reputed digital newspapers, we resorted to some of the national newspapers also. A total of 1083 comments were collected from the left-wing newspapers, along with 784 replies from the readers or other commenters. These comments were collected from 42 news articles from 7 left-wing newspapers. The politics domain contains 431 comments and 511 replies which were collected from 13 news articles from 4 English left-wing newspapers.

The sports domain consists of 318 comments and 121 replies, and the entertainment domain contains 334 comments and 152 replies. The comments and replies of the sports domain were extracted from 23 news articles from 5 newspapers, whereas the collection of the comments and replies from the entertainment domain were from 6 news articles from 3 newspapers. The overall statistics of the left-wing corpus are given in Table 5.4.

Table 5.4: Left-wing: corpus statistics

	Comments	Reply	News	Newspaper
Politics	431	511	13	4
Sports	318	121	23	5
Entertainment	334	152	6	3
Overall	1083	784	42	7

Table 5.5 depicts the statistics of the corpora with respect to the left-wing political ideology and three domains— politics, sports, and entertainment. It was observed that the Indian commenters engaged in commenting in small numbers in comparison to the UK and USA commenters. Hence, we had to use a lot of Indian news articles, to be precise 25 news articles to collect the data, whereas we collected much more comments and replies from a few news articles in the case of the UK and the USA, more specifically 6 and 11 news articles respectively. Another notable observation was that commenters from the UK were extensively engaged with themselves while commenting in comparison to their Indian and USA counterparts. This was the reason for obtaining 569 replies and 341 comments for the UK sub-corpora, whereas, the number of obtained replies was less than half and one-fourth of comments respectively for India and the USA sub-corpora.

From the right-wing English newspapers of India, the UK, and the USA, we collected a total of 951 comments along with 1220 replies. This sub-corpora was extracted from 5 leading right-wing newspapers which include ‘The Telegraph’, ‘The Daily Mail’, ‘The Times of India’, ‘The American Conservative’ and ‘Brietbart’. The overall statistics of the right-wing sub-corpora are shown in Table 5.6.

The statistics of the corpora with respect to the right-wing political ideology and

Table 5.5: Domain-wise statistics of the left-wing corpus

		Comments	Reply	News	#NP	NP Name
UK	Politics	114	399	3	1	Independent
	Sports	107	104	1	1	The Guardian
	Entertainment	120	66	2	1	The Guardian
	Overall	341	569	6	2	
India	Politics	105	104	7	2	The Hindu, The Indian Express
	Sports	107	15	17	3	The Hindu, The Indian Express, The Tribune
	Entertainment	107	19	1	1	The Indian Express
	Overall	319	138	25	3	
USA	Politics	212	8	3	1	The Huffington Post
	Sports	104	2	5	1	The Washington Post
	Entertainment	107	67	3	1	The Huffington Post
	Overall	423	77	11	2	

Table 5.6: Right-wing: corpus statistics

	Comments	Reply	News	Newspaper
Politics	327	603	10	4
Sports	311	182	8	3
Entertainment	313	435	4	3
Overall	951	1220	22	5

the three domains—politics, sports, and entertainment is represented in Table 5.7. Interestingly, we noticed that the number of comments that we collected from the UK, India, and USA sub-corpora was almost similar with a slight difference. The difference was visible in the number of replies, where it was observed that the USA commenters

Table 5.7: Domain wise statistics of the right-wing corpus

		Comments	Reply	News	#NP	NP Name
UK	Politics	105	68	3	2	The Telegraph, The Daily Mail
	Sports	103	126	1	1	The Daily Mail
	Entertainment	103	25	1	1	The Daily Mail
	Overall	311	219	5	2	
India	Politics	107	49	1	1	Times of India
	Sports	100	10	5	1	Times of India
	Entertainment	102	13	2	1	Times of India
	Overall	309	72	8	1	
USA	Politics	115	486	6	1	The American Conservative
	Sports	108	46	2	1	The American Conservative
	Entertainment	108	397	1	1	Brietbart
	Overall	331	929	9	2	

were extremely involved in the interaction with other commenters by replying to each other's comments which were almost 4 times more than the UK commenters. While on the other hand, by noticing the small number of replies in the Indian sub-corpora it can be stated that the Indian commenters were not so involved in replying back to the other commenters. A difference was noticed concerning the news articles where in the case of India and the USA, the number of news articles was 8 and 9 respectively, whereas, the UK sub-corpora was created by the comments from 5 newspapers from the UK. The creation of the UK and USA sub-corpora involves 2 English newspapers from the respective nations, while the Indian sub-corpora was curated from one Indian English newspaper.

The details of the corpus based on tokens are given in Table 5.8. These tokens were

taken into consideration while normalizing the frequencies for a specific case. For more details on normalizing this corpus, please see Chapter 6.4.

Table 5.8: Corpus statistics based on tokens

	Sports			Politics			Entertainment			Overall
	LW	RW	Total	LW	RW	Total	LW	RW	Total	
INDIA	4493	2140	6633	9247	5519	14766	3608	3768	7376	28775
UK	10037	3769	13806	8577	21373	29950	6000	2598	8598	52354
USA	5903	9280	15183	17607	22423	40030	12839	4862	17701	72914

5.4 Conclusion

In this chapter, we outlined the creation of a corpus for this doctoral dissertation. We showed the process of collecting the corpus and the problems that we faced during the corpus creation. We mentioned the reasons for selecting the particular newspapers from the three nations (UK, India, and USA). Also, we detailed how we pre-processed the corpus for annotation. We thoroughly described the corpus and depicted the statistics of the left-wing as well as right-wing sub-corpora. In the following chapter, we describe the data annotation tool and the annotation process.

Chapter 6

Method

6.1 Introduction

This chapter presents the research perspectives and the methodological approach used in the study, as well as introduces the research tool and techniques applied. This chapter begins by presenting the annotation tool that we employed for annotating the collected dataset. The following section, Section 6.3, describes the annotation process in detail. While describing the annotation process, we showed instances of different senses in identifying metadiscourse markers. Finally, in Section 6.4, we illustrate the rationale of normalization and our approach to normalizing the obtained frequencies across the corpora.

6.2 Data Annotation Tool

After the completion of the collection of data, the corpus was made ready for annotation. We used a tool named METOOL¹ to process the corpus. This tool was developed as a part of the research project *Identification and analysis of metadiscourse strategies in Spanish and in English* (IAMET), which was funded by the Ministerio de Economía

¹<https://metool.idm.upv.es/>

y Competitividad, Spain. METOOL is the outcome of this research project which was developed at the University of Wolverhampton and Universitat Politècnica de València from 2017 to 2020. Here, we are giving a brief description of the research project IAMET and METOOL.

IAMET focuses on the identification and tagging of metadiscourse strategies that are used in Spanish and in English in the genre of scientific texts, as well as the comparative analysis of the variation of these strategies in both languages. It involves three areas of knowledge: engineering, medicine, and linguistics, to determine linguistic variation when metadiscourse strategies are used among these disciplines. This project focused on the taxonomy proposed by Hyland (2005a) to classify the different elements identified. In this project, the frequencies of the elements were counted, and contrastive studies were carried out among the different disciplines written in Spanish and in English. The hypothesis involved in this project was that metadiscursive strategies are used in a different way in English and in Spanish, a fact that may influence the effectivity of communication when they are used as foreign languages. The objectives of this project were:

a) to analyze the use of metadiscourse strategies in English and in Spanish in the three different disciplines mentioned above that belong to the scientific register.

b) to study variation in the use of metadiscourse in these languages and disciplines.

In this sense, the aims of this project were twofold: first, to determine the characteristics of scientific discourse and the rhetorical strategies associated to convince readers, and second, to identify variation patterns concerning the strategies analyzed. This last aim may be of use for teachers of Spanish and English as foreign languages. The tagging and analysis were carried out with the tool 'METOOL' designed by the Research Institute for Information and Language Processing (University of Wolverhampton) for the tagging and identification of rhetorical devices in academic discourse, carried out by researchers at Universitat Politècnica de València, Universidad de Las Palmas de Gran Canaria, Universitat Jaume I and Universidad Católica de Valencia. The dif-

ferent ways of communication writers produce when persuading readers and language variation were identified in this project to show the different ways of communication of speakers of different languages and disciplines. The identification and analysis of variation in the use of rhetorical strategies in scientific papers benefit both researchers and writers of this genre. To develop this project, first, the developers compiled the corpora, second, metadiscursive strategies were identified and tagged in Spanish and in English in the three disciplines (engineering, medicine, and linguistics) and finally, metadiscourse strategies were analyzed and classified in both languages and in the three above mentioned disciplines, including examples of each category to identify the semantic implications of the elements.

This project allows scientific writers, metadiscourse researchers and teachers of Spanish and English as a foreign language to have a taxonomy, description and contextualization of metadiscourse elements that are used in English and in Spanish. Although metadiscourse strategies have been studied from different perspectives, the project proposed here that focused on the variation of metadiscourse strategies and classifies and identifies the elements in context with examples in English and Spanish, comparing three disciplines, was carried out before.

Now, we describe how METOOL works and how it detects the metadiscourse markers from written texts. The instructions for tagging a corpus with METOOL are described below:

The annotation process starts with uploading the collected dataset to the server. Then, with valid credentials, the tool could be accessed from any web browser. The following steps could be followed to annotate the corpus.

1. Click the link <https://metool.idm.upv.es/>
2. Then go to “Annotate corpus” and choose the text to annotate.
3. Click on ‘annotate’ and then choose among the three options: ‘annotate’, ‘annotate by marker’, and ‘annotate by category’. To annotate the whole text at

once, one can click ‘annotate by category’. If one wants to annotate the text by marker, then go to ‘annotate by marker’. In the following steps, we are describing how to annotate the corpus by category (annotate by category).

4. Click the option ‘annotate by category’ and go to all the categories.
5. Next, choose the first category, which is ‘Boosters’, by clicking on the option.
6. Then, the text could be seen, which is going to be annotated with the markers included in the category one has chosen highlighted in yellow.
7. After clicking the first marker that is highlighted in yellow, a menu will appear at the bottom of the screen. Please see the Figure 6.1.

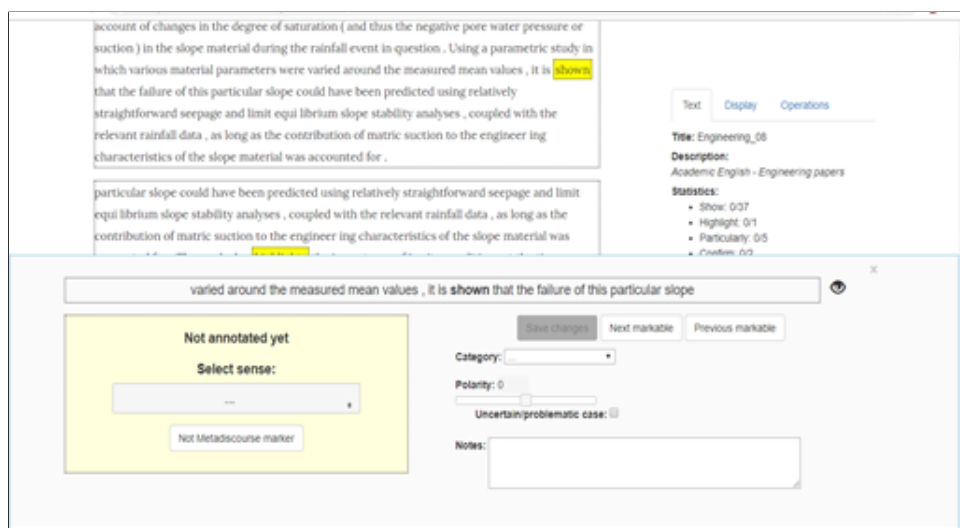


Figure 6.1: METOOL: Marker annotation

8. Next, as depicted in Figure 6.2, on top of the menu, the context of the word that is going to be tagged is in bold. Then, the sense of the marker has to be chosen from the given options if the word is considered to be a metadiscourse marker. When the sense of the word is selected automatically, one can see that the category and subcategory (if any) are inserted on the right of the menu. If the word is not a metadiscourse marker, then the option ‘not metadiscourse marker’ must be selected.

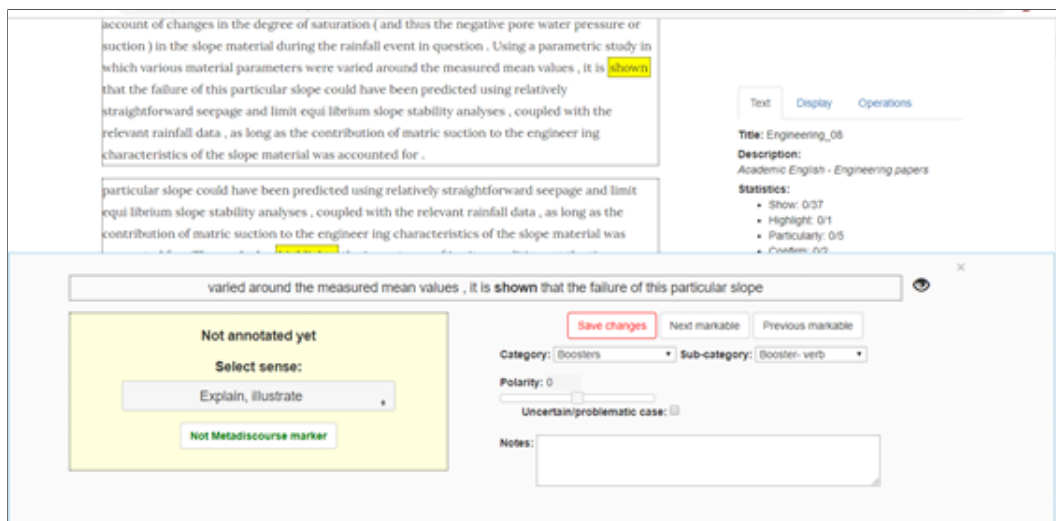


Figure 6.2: METOOL: Marker annotation

9. One can select other categories or sub-categories if necessary and can select the polarity of the marker (only some markers have polarity), for example, highly-polarity 3, extremely- polarity 5; may- polarity -3; might- polarity -5 (but this always depends on the context, of course).
10. Also, there is a column given to write any comment or remark about the chosen metadiscourse marker or to mark that this is an uncertain case. If there is nothing to write, then one has to save changes and click on the next markable.

6.3 Data Annotation Process

Initially, we uploaded the collected corpus on METOOL. Then, we followed the steps as mentioned above in Section 6.2. To ease out the annotation process, METOOL highlighted words and word phrases with probable markers and senses associated with those suggested markers. Next, we assessed the linguistic implications of the word and word phrases in the context of the comment and manually selected the most suitable marker for each word or word phrase. Also, we assessed the words or word phrases which were not detected by METOOL as markers. In Section 6.2, we described the

steps to be followed while annotating a corpus in METOOL. One can also leverage some useful functionalities of METOOL that we discuss here and which were to be followed to extract useful information from a corpus further in METOOL. METOOL provides an option named ‘Analyze Corpus’; after clicking, four further options will appear on the screen. If one has to search for an expression from the entire corpus, the first option, i.e. ‘simple search’, must be clicked. It is useful for quickly searching for an expression, a word, or a phrase. The expression that is needed to be searched must be put in the bar, and then by clicking the option ‘Analyze’, one can get the result. Another interesting functionality that the METOOL provides is for extracting concordances from the corpus. To extract the concordances from a sub-corpora, one has to click the option ‘Extract concordance from’. After clicking this option, we have to select the sub-corpora of which we want to get the concordance of the expressions (See Figure 6.3).

The search results for category *and*

The statistics and annotation are from the Left wing India politics corpus

Statistics	
Show User Marker	Sense
<input checked="" type="checkbox"/> Agnija and	It introduces additional information. Be careful, "and" is a metadisocurse marker when it joins two propositions, but not two names
<input checked="" type="checkbox"/> Agnija and	Not a marker
<input checked="" type="checkbox"/> Agnija AND	It introduces additional information. Be careful, "and" is a metadisocurse marker when it joins two propositions, but not two names
<input checked="" type="checkbox"/> Agnija And	It introduces additional information. Be careful, "and" is a metadisocurse marker when it joins two propositions, but not two names
All	None

The annotation	
1. beard , giving islamic name to children , and praise / defend those islamic terrorists	+
2. Jyotirgamaya " to know their true identi - ty , and free themselves from this DEPRESSIVE DARKNESS	+
3. follow the same people who brought death and pain to your families !! Comment - 3 : MUGAL	+
4. did not use the language , come brothers and sisters join us They used	+
5. They used rabid violence and most demeaning language . The choice was	+
6. language . The choice was to either fight and may die , or accept SLAVERY , conversion	+
7. conversion was NOT done to create " Muslims " and give them EQUAL RIGHTS / RESPECT ! , it was	+
8. , kill your children , rape your women , and then ask you to join their gang ?? Comment	+

Figure 6.3: METOOL: Concordance

Another useful functionality of METOOL is that it helps to extract the statistics per category from the corpus. Clicking the option ‘Extract statistics and search corpus by’, one has to select a sub-corpora. Then the statistics of the markers per category

Statistics per categories

Corpus: Left wing UK politics

Boosters: 74	booster- noun: 20
	Booster- verb: 26
	Booster- adjectives: 10
	Booster- adverbs: 16
	Booster- phrases: 2
Logical markers: 221	Additive: 114
	Contrastive: 59
	Consecutive: 48
Code Glosses: 24	Reformulation: 13
	Exemplification: 11
Frame markers: 60	Sequencers: 15
	Topicalisers: 2
	Discourse labels: 0
	Announcers: 43

Annotated markers from **the Right wing USA politics corpus**

Punctuation

:: 263

A

about: 98

above: 4

acceptable: 2

actually: 29

aim to: 1

almost: 6

also: 14

always: 18

analysis: 1

and: 617

any: 48

anyone: 5

appear to: 2

(a) Statistics per category of a sub-corpora (b) Report from a sub-corpora after annotation

Figure 6.4: Analysis of Corpus: snapshots from METOOL

are shown (See Figure 6.4a).

Also, METOOL helps to report the annotated markers from the selected sub-corpora. One has to click the option, ‘Report from a corpus’, and select the sub-corpora, and one can get the report of the annotated markers as in Figure 6.4b. Moreover, METOOL provides the option of getting the quantitative results in Excel format, which is extremely convenient. We leverage all these useful functionalities of METOOL in our annotation process.

6.3.1 Metadiscourse Markers with Different Senses

Identifying the appropriate sense of a word in the given context is crucial for identifying whether a word will be a metadiscourse marker or not. Here, we are giving some comments as examples that are taken from our corpus.

The given examples illustrate the different senses of *you* that we captured from different comments.

Example- (a): *So you are happy to use misinformation re car sales? It was a record year for new car sales last year, that is partly why sales are down.* (Corpus:

UK-POL:1)

METOOOL: Engagement Markers: Personal references

Example- (b): *So what May is saying to parliament is, ‘we have a secret dossier which **you** are not allowed to see or scrutinise, but trust us it says that Russia did it’. Does she really expect that to work after Iraq?* (Corpus: UK-POL:1)

METOOOL: Engagement Markers: Personal references.

Justification: Here, in both cases, ‘you’ is a metadiscourse marker with different senses. In the first comment Example- (a), by saying ‘you’ the person who is replying is commenting on the commenter’s response. In the second comment Example- (b), by saying the word ‘you’, the commenter is referring to something that Teresa May (former Prime Minister of the UK) said. Therefore, in both the example comments, ‘You’ is a metadiscourse marker when the writer is referring to the reader or the writer is indicating evidentiality.

6.4 Normalization

As the length (i.e., word counts) of the different corpora are not the same, the three datasets are not comparable. Therefore, to compare the datasets, ‘normalization’ was a crucial step to enable the correct comparison of corpora of different lengths. ‘Normalization’ is a way to convert raw counts into rates of occurrences, which enables the scores from texts of different lengths can be compared (Lüdeling and Kytö, 2008). Normalization is very useful in linguistic studies considering the fact that it makes us enable to compare linguistic features of corpora of distinct lengths (Friginal and Hardy, 2013).

In the literature, Biber et al. (1998) proposed a convenient formula for normalizing frequencies across corpora. According to them, the raw frequency count should be divided by the number of words in the corpus and then multiplied by whatever basis is chosen for norming. Following Blagojević (2016), we normalized each marker’s

frequency count as follows:

$$\text{Normed rate of occurrence} = \frac{\text{Raw frequency of a feature in corpus}}{\text{Total number of words in corpus}} \times \text{Norming number}$$

In the present study, the frequency count of each of the markers is normed to a basis of 10,000 words.

6.5 Conclusion

This chapter outlines the method used in this doctoral dissertation. It provides an overview of METOOL, a web-based annotation tool for tagging and identification of rhetorical devices in academic discourse, that we employed to annotate our dataset. Further, we provide details of the data annotation process. We also discuss the normalization for quantitative comparisons of our cross-cultural datasets with different sizes. The next chapter describes the results in a detailed way that includes the quantitative analysis of the news comments based on country, domain, political ideology, and cross-culture. Moreover, we present an in-depth analysis of the occurrences of each category of metadiscourse markers.

Chapter 7

Results & Discussions

7.1 Introduction

This chapter presents the findings of various investigations undertaken using metadiscourse markers by the commenters of three nations- the United Kingdom, India, and the United States of America. We elicit the results based on the use of normalized¹ frequencies that are estimated with respect to different factors—country, domain, political ideology, English native and non-native speakers.

The rest of the chapter is organized as follows: in Section 7.2, we present the quantitative analysis of India comments; the following two sections describe the quantitative analysis of the UK (Section 7.3) and the USA (Section 7.4) comments. In Section 7.5 we further dig into the in-depth quantitative and qualitative analysis of the obtained markers; Section 7.6 demonstrates the quantitative analysis of employing metadiscourse markers across three different prior selected domains; in Section 7.7, we provide the quantitative analysis of using metadiscourse markers with respect to political ideology; and Section 7.8 demonstrates the quantitative analysis of using metadiscourse markers in the context of cross-culture. Finally, we conclude this chapter in Section 7.9.

¹All the reported frequencies are normalized (cf. Section 6.4) per 10,000 words.

7.2 Quantitative Analysis of India Comments

We collected a total of 838 digital news comments by the India commenters. After analyzing the comments, we found that per 10,000 tokens, there were 5604.74 metadiscourse markers in the India sub-corpora, among which 1852.19 were *interactive* markers, and 3752.93 were *interactional* markers. We reported the overall quantitative usage of *interactive* and *interactional markers* by the India commenters in Table A.1 and Table A.2 (See Appendix). The distribution of *interactive* and *interactional markers* for India commenters is depicted in Figure 7.1 and Figure 7.2.

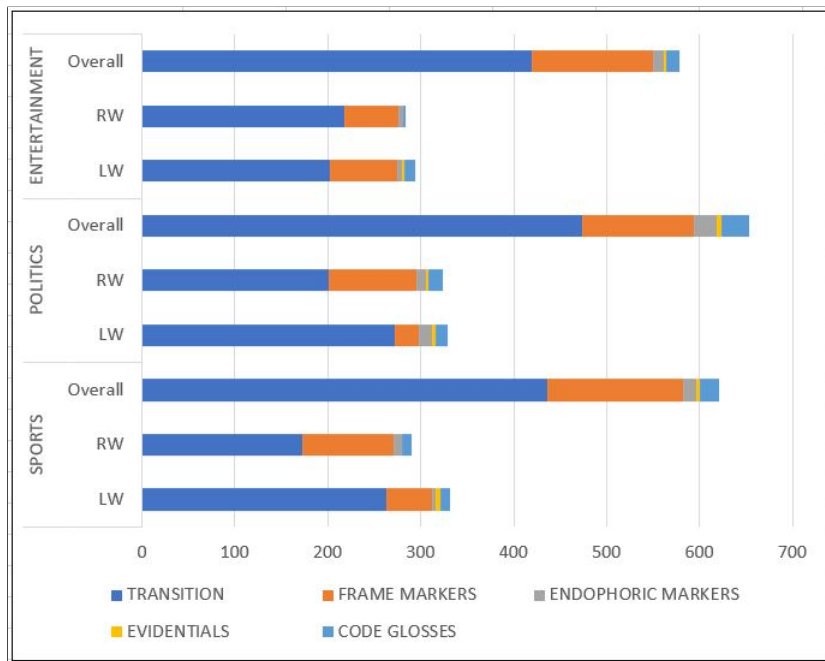


Figure 7.1: India: Distribution of *Interactive markers* (normalized)

Irrespective of domains, while using the interactive markers, the India commenters tended to use *transition* markers the most. In the overall use of *transition markers*, the commenters of the politics domain used them more to make the readers (in our case, other commenters) understand how the commenters are linking the propositions in an argument, followed by the sports and entertainment domains. Following that, quantitatively, they used *frame markers*. Noticeably, in the case of *endophoric markers*, the overall use was seen more in the politics domain, followed by the sports and

entertainment domains. This suggests that the commenters of the politics domain are more prone to convey to their readers (in our case, other commenters) the semantic relations between propositions. The overall use of *evidentials* was seen to be less in all three domains (sports, politics, and entertainment) compared to the other interactive markers. In the entertainment domain, *evidentials* were used 2.77 per 10,000 words which is comparatively less compared to the sports (4.45 per 10,000 words) and politics domain (5.06 per 10,000 words). *Code glosses* were used most frequently in the politics domain, followed by the sports and entertainment domain. So, in the overall use of the interactive markers, it was noticed that except *frame markers* (used most frequently in the sports domain followed by the entertainment and politics domains), all the other interactive markers were used most frequently in the politics domain followed by the sports and entertainment domains as the commenters felt the need to guide their readers.

Considering political ideology, we noticed that in the sports domain, *transition markers* and *code glosses* were used more frequently by the left-wing commenters than the right-wing. On the contrary, the right-wing commenters used *frame markers* and *endophoric markers* more than the left-wing. A possible reason for this difference could be that left-wing commenters tend to let their readers know the linking and the shifting of the propositions in an argument. However, right-wing commenters have the habit of explaining the situations by rephrasing or reformulating the main propositions of an argument. In the use of *evidentials*, it was observed that the right-wing commenters of the sports domain never used *evidentials* for a single time; however, they were used 4.45 per 10,000 words by the left-wing. In the politics domain, the left-wing commenters used *transition markers*, *endophoric markers*, and *evidentials* more, whereas the right-wing commenters were prone to use *frame markers* and *code glosses* more. In the entertainment domain, the presence of *transition markers* was more in the right-wing sub-corpora, whereas the *frame markers* and *code glosses* were used more in the left-wing sub-corpora. *Endophoric markers* were used almost the same in both left-wing

and right-wing sub-corpora which suggests the fact that the left-wing and right-wing commenters felt the need to refer to the other parts of the comments for the better understanding of the commenters. There was not a single use of *evidentials* in the right-wing sub-corpora of the entertainment domain; however, the left-wing commenters used them very less (2.77 per 10,000 words).

The quantitative analysis of interactional metadiscourse markers for India commenters is revealed in Table A.2 and Figure 7.2.

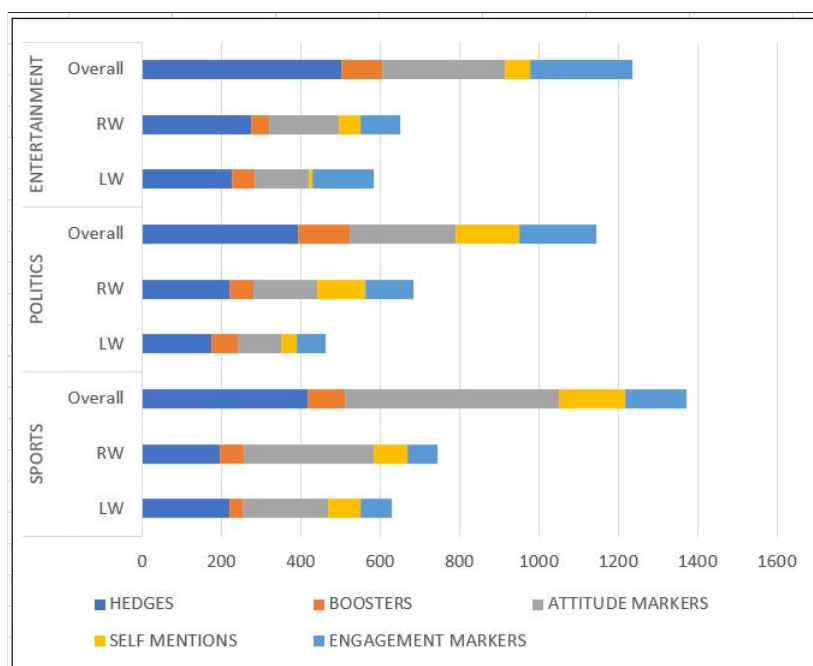


Figure 7.2: India: Distribution of *Interactional markers* (normalized)

If we consider the domain wise use of interactional markers, commenters from the entertainment domain employed *hedges* more frequently followed by the sports and politics domain. In the use of *boosters*, the commenters of the politics domain used them more, followed by the entertainment and sports domains. This shows that the commenters of the entertainment domain tend to restrain themselves from making any strong commitments. Rather they subside from the situation in a polite manner. On the other hand, the commenters of the politics domain were certain about their commitments. So, they use *boosters* to emphasize the certainty of the commitments.

Attitude markers were used most frequently in the sports domain, followed by the entertainment and politics domain. *Self-mentions* were used almost the same with a slight difference in the sports and politics domain (166.46 and 162.14 per 10,000 words, respectively); however, they were used comparatively less in the entertainment domain (64.17 per 10,000 words). In the use of *engagement markers*, the commenters of the entertainment domain used them more to include the other commenters in the conversation, followed by the politics and sports domains.

With respect to political ideology, we can see that in the sports domain, *hedges* and *engagement markers* were used comparatively more in the left-wing sub-corpora, whereas *boosters*, *attitude markers*, and *self-mentions* were used more frequently in the right-wing sub-corpora. This implies that the left-wing commenters of the sports domain have a tendency to include the readers (in our case, the other commenters) in the conversation by stating his/her view. Contrarily, the right-wing commenters of the sports domain were more inclined to show the readers their certainty and attitude regarding a particular proposition. In the use of *self-mentions*, it was noticed that in both the sub-corpora, the use was almost the same with a slight difference. A potential reason could be that the India commenters were more inclined to proclaim the authorial self while commenting on digital news. In both the politics and entertainment domains, *hedges*, *attitude markers*, *self-mentions*, and *engagement markers* were used comparatively more in the right-wing sub-corpora than the left-wing, whereas *boosters* were used slightly more in the left-wing sub-corpora than the right-wing.

A comparative analysis of using *interactive* and *interactional* metadiscourse markers by the India commenters on the basis of Table A.1 and Table A.2, clearly shows the followings:

- All the *interactive* and *interactional* markers were used except *evidentials*, which was not used for a single time in the right-wing sub-corpora of the sports domain and left-wing sub-corpora of the entertainment domain. This hints at the tendency of the India commenters to guide the readers (in our case, the other

commenters) and also involve them in the arguments.

- The India commenters tend to use *interactional markers* more extensively than the *interactive markers* across all the domains (i.e., sports, politics, and entertainment) and political ideology (i.e., left-wing and right-wing). More precisely, they used *interactional markers* almost twice or more than twice than they used *interactive markers* across domains. This highlights the inclusive nature of the India commenters who want to involve the other commenters in the conversation.
- Taking into account the use of *interactive markers* based on political ideology, it can be stated from Figure 7.1 that the overall use of *interactive markers* was seen more in the left-wing sub-corpora than the right-wing which suggests the left-wing commenters of India want to guide the readers through the arguments though they are not so keen to include them.
- From Figure 7.2, it is evident that in all the three domains, the *interactional markers* were used comparatively more in the right-wing sub-corpora than the left-wing.
- The *interactive markers* were used most by the India commenters of the politics domain, followed by the sports and entertainment domains, whereas the *interactional markers* were used mostly in the sports domain, followed by the entertainment and politics domain.

7.3 Quantitative Analysis of UK Comments

For the UK sub-corpora, we collected a total of 1440 comments. We found that per 10,000 tokens, there were 6331.75 metadiscourse markers, among which 2133.06 were *interactive markers* and 4198.69 were *interactional markers*. The overall usage of *interactive* and *interactional* metadiscourse markers by the UK commenters is portrayed in

A.3 and A.4 (See Appendix). The distribution of *interactive* and *interactional markers* for the UK commenters is depicted in Figure 7.3 and Figure 7.4.

While using the *interactive* markers, the UK commenters used *transition markers* extensively across domains, in particular, almost equally in the entertainment and sports domain and a bit less in the politics domain.

Next, they used *frame markers*; however, the use was way behind *transition markers*. Specifically, the commenters of the politics domain used *frame markers* most (148.56 per 10,000 words) compared to the other two domains — 80.32 and 64.13 per 10,000 words in sport and entertainment, respectively which indicates that the commenters were eager to let their readers (in our case, the other commenters) know the sequences and different stages of the conversation for better understanding. In terms of usage, the UK commenters used the other interactive markers in the following order — *code glosses*, *endophoric markers*, and *evidentials*. We noticed that almost all the *interactive markers* were used across all the domains by the UK commenters except *evidentials*, as in the right-wing newspapers of the sports domain and in the left-wing newspapers of the entertainment domain, there was not a single use of *evidentials* (See Table A.3). A possible reason for not using *evidentials* by the commenters of these particular domains and political ideology could be that they were not interested in referring to the authentic source of their argument.

Figure 7.3 depicts that the interactive markers were mostly used in the left-wing sub-corpora in comparison to the right-wing. However, irrespective of domains, the use of *evidentials* was much less, in particular, 4.98 per 10,000 words in the sports domain, 8.41 per 10,000 words in the politics domain, and 3.85 per 10,000 words in the entertainment domain, respectively. Again, as a possible reason, it could be assumed that the commenters of the three domains did not feel the need to show their readers the authenticity of their argument by referring to its source information.

Considering the domain wise use of *interactional markers*, we observed that the UK commenters tended to use *hedges* more often, followed by *attitude markers*, *engagement*



Figure 7.3: UK: Distribution of *Interactive markers* (normalized)

markers, *boosters*, and *self-mentions* in sports, politics, and entertainment domains. The significant use of *hedges* by the commenters indicates the commenter's disinterestedness in showing the readers (in our case, the other commenters) full commitment to statements. In the use of *self-mentions*, we observed that in the entertainment and sports domains, they were used significantly less (20.57 and 13.21 per 10,000 words, respectively) compared to the politics domain (106.90 per 10,000 words). Figure 7.4 reveals that in the politics domain, the *interactional markers* were mostly used by the UK commenters, followed by the sports and entertainment domains which suggests that the commenters were eager to involve the readers in the argument.

Considering the political ideology wise use of the *interactional markers*, we noticed that, in the sports domain, the left-wing commenters used *attitude markers* and *engagement markers* more, whereas *hedges*, *boosters*, and *self-mentions* were used more frequently in the right-wing sub-corpora. This means the left-wing commenters perceived the need to engage the readers (in our case, the other commenters) in the

conversation more. Interestingly, some of the right-wing commenters, on the one hand, showed uncertainty about the commitments by using *hedges*, and on the other hand, some showed their full commitment to propositions by using *boosters* more. In the case of the politics domain, we found that *hedges*, *attitude markers*, and *engagement markers* were used comparatively more in the left-wing sub-corpora, while *boosters* and *self-mentions* were used more frequently in the right-wing sub-corpora. *Boosters* were used almost the same in the politics domain's left-wing and right-wing sub-corpora. *Self-mentions* were used enormously in the right-wing sub-corpora (103.40 per 10,000 words) than the left-wing (3.50 per 10,000 words) which suggests that the right-wing commenters were keener to include their authorial self in the conversation than the left-wing ones. The left-wing commenters of the entertainment domain used *hedges* and *boosters* (a bit more) more than the left-wing, while the right-wing commenters used *attitude markers*, *self-mentions*, and *engagement markers* comparatively more than the left-wing. The use of interactional markers by the right-wing commenters implies the following three things: first, the commenters of the entertainment domain were eager to show their attitude towards the propositions of an argument; second, they mention the authorial self in the conversation to create an alliance with the readers (in our case, the other commenters), and third, they want to build a relationship with the readers (the other commenters) by engaging them more in the argument. *Self-mentions* were used significantly less in the left-wing sub-corpora (1.67 per 10,000 words) than in the right-wing (11.55 per 10,000 words), which indicates the unwillingness of the commenters to refer to the authorial self in the argument.

If we conduct a comparative analysis of the *interactive* and *interactional markers* by the UK commenters based on Table A.3 and Table A.4, and Figure 7.3 and Figure 7.4, the findings are as follows:

- It is noted that the *interactional markers* were used significantly more by the UK commenters compared to the *interactive markers* (quantitatively even twice or more than twice). The tendency to involve the readers (in our case, the other

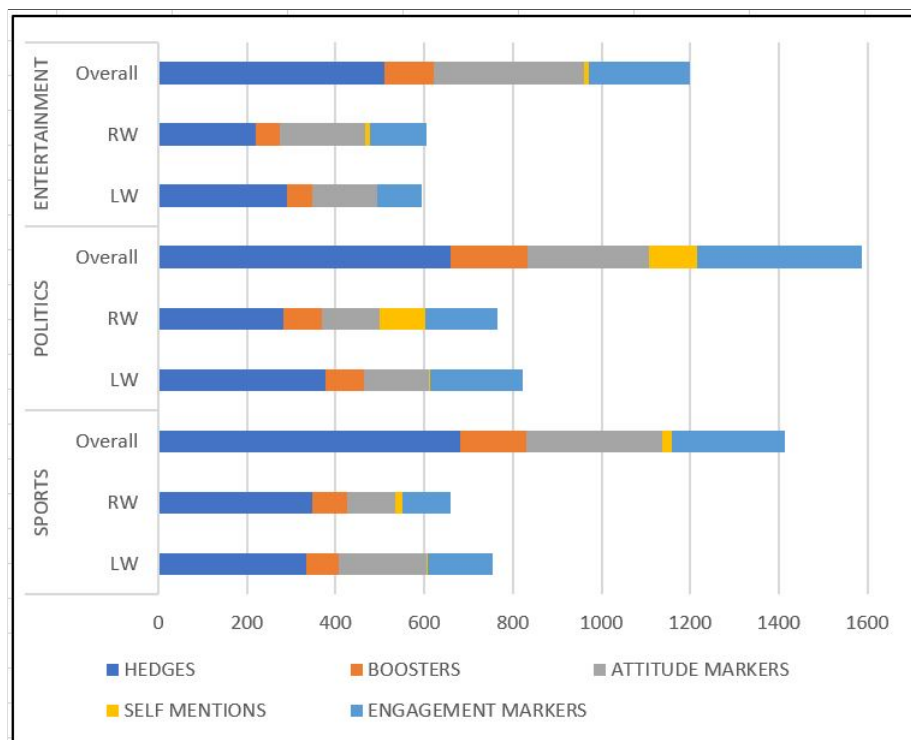


Figure 7.4: UK: Distribution of *Interactional markers* (normalized)

commenters) in the conversation might be one possible reason for the significant use of the *interactional markers*.

- All the *interactive* and *interactional markers* were used irrespective of domains and political ideology, except for *evidentials* (from *interactive markers*), which is not used for a single time in the right-wing sub-corpora of sports domain and the left-wing sub-corpora of entertainment domain.
- The *interactive markers* were used more in the left-wing sub-corpora irrespective of all domains (sports, politics, and entertainment) which suggests that the commenters of the left-wing sub-corpora understand the need to guide the readers (in our case, the other commenters) through the text to comprehend the progression of the argument.
- For *interactional markers*, in the sports and politics domain, the commenters of left-wing newspapers used more *interactional markers* than the right-wing

newspapers. However, for the entertainment domain, the right-wing commenters used more *interactional markers* than the left-wing (See figure 7.4).

- The commenters of the politics domain used both the *interactive* and *interactional* markers more, followed by the sports and entertainment domains (See Figure 7.3 and Figure 7.4).

7.4 Quantitative Analysis of USA Comments

We collected a total number of 1760 comments for the USA sub-corpora. After analyzing the corpus, we found per 10,000 tokens, 5278.6 metadiscourse markers were used by the commenters, among which 1835.27 were *interactive* markers and 3443.33 were *interactional* markers. We detailed the overall usage of using *interactive* and *interactional* metadiscourse markers by the USA commenters in Table A.5 and Table A.6 (See Appendix). The distribution of *interactive* and *interactional* markers used by the USA commenters are portrayed in Figure 7.5 and Figure 7.6.

In the sports domain, *transition markers* and *frame markers* were used significantly more compared to the other two domains. *Endophoric markers* and *evidentials* were used less irrespective of all the domains. *Endophoric markers* were used almost the same (with a slight difference) in the sports and politics domain (21.40 and 19.47 per 10,000 words, respectively), whereas, in the entertainment domain, they were used only by the left-wing commenters (4.67 per 10,000 words). One probable reason for this could be that the commenters of the entertainment domain were reluctant to refer to information from the other parts of the conversation. On the other hand, *evidentials* were used almost similarly in the politics and entertainment domain (7.67 and 6.73 per 10,000 words, respectively) and comparatively more in the sports domain (10.93 per 10,000 words). *Code glosses* were used more in the sports domain, followed by the politics and entertainment domains which shows the commenters' intention to make things more understandable for the readers (in our case, the other commenters) by

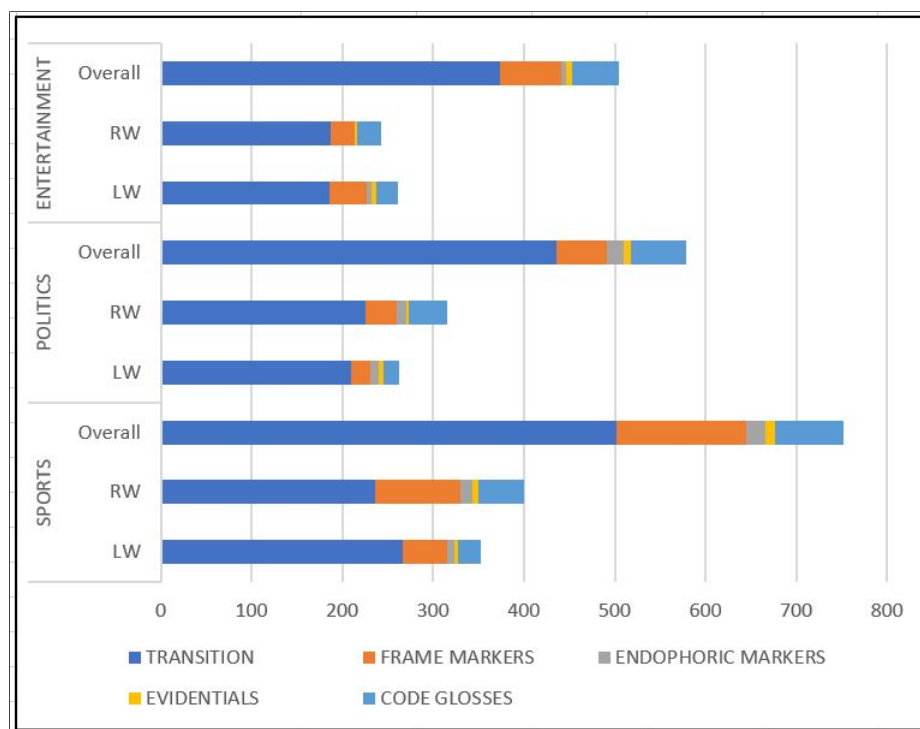


Figure 7.5: USA: Distribution of *Interactive markers* (normalized)

reformulating a particular idea or proposition..

Regarding the use of *interactive* markers based on political ideology by the USA commenters, it was seen that in the sports domain, only *transition markers* were used comparatively more in the left-wing sub-corpora, whereas the rest of the interactive markers (*frame markers*, *endophoric markers*, *evidentials*, and *code glosses*) were used more in the right-wing sub-corpora. As a reason for that, it could be assumed that the left-wing commenters of the sports domain were eager to let their readers (in our case, the other commenters) know the semantic relation between the ideas or propositions in an argument. In the politics domain, except *evidentials*, all the other interactive markers (*transition markers*, *frame markers*, *endophoric markers*, and *code glosses*) were used more in the right-wing sub-corpora, whereas *evidentials* were used comparatively more in the left-wing sub-corpora. In the entertainment domain, *transition markers* and *code glosses* were used comparatively more in the right-wing sub-corpora, while *frame markers* and *evidentials* were used more frequently in the left-wing sub-corpora.

In the case of *endophoric markers*, there was not a single use in the right-wing, whereas it was seen to be used less in the left-wing sub-corpora (4.67 per 10,000 words) which suggests that the right-wing commenters of the entertainment domain were not so eager to refer to information from the other comments, rather they want to maintain the flow of the casual conversation.

We observed that while using the *interactional markers*, the USA commenters tended to use *hedges* extensively across all the domains which implies the commenters' reluctance to show full certainty to a commitment or a proposition. *Boosters* were used comparatively less in the sports domain and almost more than twice in the politics and entertainment domains. In the case of *self-mentions*, the commenters of the sports and politics domain used almost the same, whereas in the entertainment domain, they were used a bit less. *Engagement markers* were used almost the same in the sports and entertainment domains, while they were used a bit less in the politics domain (See Figure 7.6). Considering the domain wise use by the USA commenters, *interactional markers* were mostly used in the sports domain, followed by the politics and entertainment domains. It was also observed that irrespective of domains, *interactional markers* were used mostly in the comments of the right-wing sub-corpora to include the other commenters in the conversation. (See Figure 7.6).

With respect to political ideology, it was seen that in the sports domain, *hedges*, *attitude markers*, and *self-mentions* were used more frequently by the right-wing, whereas *boosters* and *engagement markers* were used more by the left-wing commenters. A significant difference was noticed in the use of *boosters*, where the right-wing commenters of the sports domain used them very less (14.01 per 10,000 words) than the left-wing (79.62 per 10,000 words), which was almost 5 times more. It can be assumed that the left-wing commenters of the sports domain were more willing to emphasize certainty to a proposition in an argument. In the politics domain, all the *interactional markers* were used comparatively more in the right-wing sub-corpora than the left-wing. In the entertainment domain, *hedges*, *boosters*, and *attitude markers* were used more in

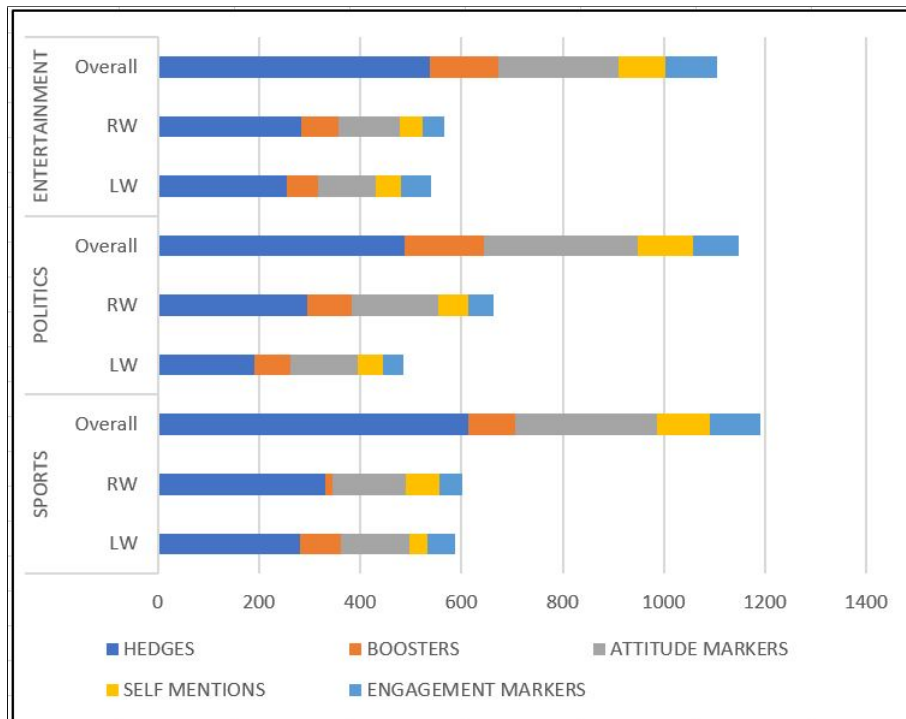


Figure 7.6: USA: Distribution of *Interactional markers* (normalized)

the right-wing, while *self-mentions* and *engagement markers* were used more in the left-wing sub-corpora.

A comparative analysis of the use of *interactive* and *interactional* markers by the USA commenters based on Figure 7.5 and 7.6 represents the following facts:

- Across all the domains, all the *interactive* and *interactional* markers were used except for the entertainment domain where there was not a single use of *endophoric markers* (a lexico-grammatical category of *interactive* markers) in the comments of the right-wing sub-corpora.
- Overall, *interactional* markers were used most frequently in comparison to the *interactive* markers, which is more than 1.5 times. This suggests the commenters' intention of involving their readers (in our case, the other commenters) in the conversation.
- Irrespective of domains, *interactional* markers were mostly used in the comments

of the right-wing sub-corpora compared to the left-wing ones by the USA commenters (See 7.6).

- Figure 7.5 illustrates the fact that in the comments of the right-wing sub-corpora, *interactive* markers were mostly used compared to the left-wing except in the entertainment domain, where in the comments of left-wing sub-corpora *interactive* markers were mostly used. Irrespective of domains and political ideology, *evidentials* were used minimally, indicating the commenters' intention to converse casually.
- If we compare the domain wise use of the markers, we can see that both the *interactive* and *interactional* markers were used more in the sports domain, followed by the politics and entertainment domains as the commenters of the sports domain focus on guiding their readers (in our case, the other commenters) as well as including them in the conversation.

7.5 In-depth Analysis of Markers

In this section, we provide the in-depth analysis of the metadiscourse markers (*interactional* and *interactive*) that were employed by the UK, India, and the USA commenters. This discussion further digs into the characteristics of the markers, which helps to capture the potential differences in the expressive nature of the commenters. This, in turn, could be very useful for comparing the expressive nature of different groups of commenters. In the subsequent sections, we present the further quantitative details of the *interactional* and *interactive* markers noted in our corpus. We also present the comparisons of the occurrences of *interactional* and *interactive* markers based on their lexico-grammatical categories. It would be worth investigating to analyze the lexico-grammatical categories further. Hence, we extracted the items which were tagged for a specific lexico-grammatical category of a marker. We dig into the detailed analysis of each of the *interactional* and *interactive* markers. However, accommodating every

occurrence of the lexico-grammatical categories in a table is cumbersome. Hence, we provide the most frequent words of the lexico-grammatical categories.

Also, here we discuss with examples from each country along with the sense in which it is used in the comment by the commenters. Also, we would like to mention here that while demonstrating the examples, we mentioned the source of that particular example, for instance, left-wing (LW) or right-wing (RW) for political ideology, for a specific country, i.e. the United Kingdom (UK), The United States of America (USA), and India (IND), and for specific domains, entertainment (ENT), politics (POL), and sports (SPR). However, as we have mentioned before that we have presented the occurrences without segregating the different forms of the same lexeme; the examples presented here from our corpus include the different forms of the same lexeme.

7.5.1 Interactional: Hedges

According to Hyland (2010, p. 129), *hedges* are used by the writers to ‘withhold writer’s full commitment to proposition’. It is one of the crucial interactional metadiscourse categories often used by the writers to withhold full responsibility and certainty to propositions. The writers use *hedges* to express the possibility, uncertainty, and negativity of a proposition, and by using *hedges*, the writer conveys politeness and ambiguity to persuade the readers. They also project the unwillingness of the writer to indicate propositional information. Previously, various authors such as Alonso et al. (2012), Crompton (2012), Hyland (1995), Hyland (1996) have studied *hedges* where they were used by the writers to express uncertainty towards a proposition.

Specifically in editorials, articles, and news reports published in digital newspapers, Biri (2018), Hooi et al. (2020), Wang et al. (2016), Carrió-Pastor and Alonso-Almeida (2019), studied *hedges* which were used by the writers to withhold commitment to a statement. However, the editorials and articles published in digital newspapers differ from digital news comments because editorials and articles are bound to be formal in nature, whereas digital news comments are not bound to be formal and usually infor-

mal in nature. Nevertheless, this informal nature does not bind the writers to be polite. In our study, *hedges* were used by the commenters in a five-fold way– 1) to indicate the commenter’s degree of uncertainty or confidence about statements, 2) to express the commenter’s personal beliefs and doubts, and subjective opinion, 3) to express politeness and vagueness, 4) to propose suggestions and alternative solutions, 5) to withhold commitments and responsibility to propositions and act more as an opinion rather than as a fact. In our study, *hedges* were used extensively by the commenters that inline with the studies of Mur-Dueñas (2011), Hyland (2005a), and Hyland (2005b). There are various taxonomies of *hedges* that different authors proposed, such as Mur-Dueñas (2011), Hyland (2005a), Crompton (2012), and our study followed the categorization of *hedges* by these authors where *hedges* are divided into six lexico-grammatical categories of *hedges*: verbs, nouns, modal verbs, adverbs, adjectives, and phrases. In our corpus, modal verbs were used most frequently, and nouns and phrases were used less frequently by the commenters of the UK, India, and the USA.

Table 7.1: Normalized frequencies of *hedges* found in India, the UK and the USA

	Entertainment			Politics			Sports			
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall	
INDIA	Nouns	2.77	0.00	2.77	5.41	1.81	7.22	2.23	0.00	2.23
	Verbs	66.52	84.93	151.44	19.47	48.92	68.39	42.29	46.73	89.02
	Modal verbs	97.01	90.23	187.24	90.84	96.03	186.87	97.93	84.11	182.04
	Adjectives	13.86	39.81	53.67	20.55	32.61	53.16	44.51	42.06	86.57
	Adverbs	36.03	55.73	91.76	32.44	38.05	70.49	33.39	14.02	47.40
	Phrases	11.09	2.65	13.74	4.33	1.81	6.14	0.00	9.35	9.35
	Total	227.27	273.35	500.63	173.03	219.24	392.27	220.34	196.26	416.60
UK	Nouns	11.67	0.00	11.67	16.32	24.33	40.65	14.94	5.31	20.25
	Verbs	73.33	46.19	119.52	79.28	54.27	133.56	68.75	76.94	145.69
	Modal verbs	60.00	76.98	136.98	158.56	100.59	259.16	121.55	108.78	230.33
	Adjectives	90.00	73.13	163.13	68.79	50.06	118.85	74.72	103.48	178.20
	Adverbs	33.33	15.40	48.73	40.81	43.51	84.32	44.83	39.80	84.63
	Phrases	21.67	7.70	29.36	12.82	8.89	21.71	8.97	13.27	22.23
	Total	290.00	219.40	509.40	376.59	281.66	658.25	333.77	347.57	681.34
USA	Nouns	4.67	6.17	10.84	10.79	16.95	27.74	13.55	12.93	26.48
	Verbs	53.74	65.82	119.56	43.16	58.42	101.59	54.21	54.96	109.17
	Modal verbs	91.13	90.50	181.63	57.93	98.11	156.05	96.56	54.96	151.52
	Adjectives	62.31	59.65	121.96	39.76	51.29	91.04	57.60	72.20	129.80
	Adverbs	36.61	53.48	90.08	35.78	58.87	94.65	45.74	68.97	114.70
	Phrases	6.23	8.23	14.46	3.98	12.49	16.46	13.55	10.78	24.33
	Total	254.69	283.83	538.53	191.40	296.12	487.53	281.21	274.78	556.00

Table 7.1 represents the normalized frequency of *hedges* found in the UK, India and the USA sub-corpora. Here, we present our analysis based on domains and political

ideology.

The commenters from the three nations (the UK, India and the USA) used the following *hedges* in their comments: verbs (*want, think, seem, suggest, feel, argue, expect, regard, assume, attempt, consider, predict, see, propose, imply, tend, believe, indicate, presume, point to, interpret, posit, appear*), nouns (*idea, perspective, view, belief, argument, possibility, proposition, probability, tendency, prediction, expectation, interpretation, notion, hypothesis*), adverbs (*just, probably, quite, perhaps, nearly, maybe, somewhat, slightly, presumably, about, typically, often, almost, partly, basically, unlikely, potentially, sometime, likely, normally, approximately, commonly, virtually, usually, roughly, not necessarily, relatively*), adjectives (*some, common, certain, hypothesize, potential, anyone, any, likely, certain amount, possible, typical, common, propose, unclear, feasible, indicative*), phrases (*in theory, like, in part, at least, in general*) and modal verbs (*can, may, shall, will*).

India: The quantitative analysis of *hedges* used by the India commenters considering different domains reveals that the Indians used *hedges* more in the entertainment domain, followed by the sports and politics domain which suggests the tendency of the commenters of the entertainment domain to show their uncertainty to a commitment or a particular proposition. In the entertainment domain, with respect to usage frequencies, the Indians used the lexico-grammatical categories in the following order: modal verbs, verbs, adverbs, adjectives, phrases, and nouns. Regarding political ideology, it was noticed that the right-wing commenters (273.35 per 10,000 words) used overall hedges more than the left-wing commenters (227.27 per 10,000 words). It can be seen from Table 7.1 that although in overall use of *hedges* right-wing precedes, in the case of hedges-nouns there was not a single use by the right-wing commenters while the left-wing commenters used it for 2.77 per 10,000 words. Significant differences were noticed in the use of adjectives and phrases. The left-wing commenters used adjectives much less (for 13.86 per 10,000 words), whereas the right-wing commenters used them comparatively more (for 39.81 per 10,000 words) which shows the tendency of the com-

menters' to use attribution to the nouns while showing their uncertainty regarding a particular proposition. However, in the use of phrases, it was seen that the left-wing commenters used phrases much more (11.09 per 10,000 words) than the right-wing commenters (2.65 per 10,000 words). In the sports domain, the overall usage of the lexico-grammatical categories of *hedges* was found more on the left-wing compared to the right-wing. Likewise the entertainment domain, the same thing was observed in the use of nouns, where not a single noun was used by the right-wing commenters, whereas the left-wing commenters used it for 2.23 per 10,000 words. The verb, modal verb, and adjective usage are almost the same in both left and right-wing, with a slight variation. Nevertheless, it was noticed that the left-wing commenters used adverbs significantly more (33.39 per 10,000 words) than the right-wing commenters (14.02 per 10,000 words). Also, hedges-phrases were not seen to be used in the left-wing corpus, although the right-wing commenters used it for 9.35 per 10,000 words. In politics, the overall usage of the lexico-grammatical categories was seen more by the right-wing commenters than the left-wing. The difference was significant in the use of nouns and phrases, where the left-wing commenters used more than the right-wing commenters. On the contrary, in using verbs, modal verbs, and adjectives, the right-wing commenters used more than the left-wing commenters of the India sub-corpora.

The UK: Table 7.1 portrays that the UK commenters of the sports domain used the lexico-grammatical categories of *hedges* significantly more than the politics and entertainment domain. Concerning political ideology, we found that the overall usage of the lexico-grammatical categories of *hedges* by the right-wing and left-wing commenters was almost the same. Probably, the commenters of both the right-wing and left-wing sub-corpora of the UK were more eager to imply subjectivity through their opinions and, at the same time, shunning certainty to a particular proposition. Regarding the frequency of using verbs, modal verbs, adjectives, adverbs, and phrases, it was observed that it was almost similar in the left-wing and right-wing sub-corpora. However, in the use of nouns, a difference was noticed (although the number is on the lower side)

where the right-wing commenters of the sports domain used nouns for 5.31 per 10,000 words, whereas the left-wing commenters used them much more, 14.94 per 10,000 words (almost three times). In the politics domain of the UK, the left-wing commenters used the lexico-grammatical categories (verbs, modal verbs, adjectives, and phrases) more than the right-wing. On the other hand, the right-wing commenters used nouns and adverbs more than the left-wing. The left-wing commenters used modal verbs almost 1.5 times more than the right-wing which hints at the fact that the left-wing commenters of the UK express their thoughts in an indirect way to the readers (in our case, the other commenters). In the entertainment domain, the overall usage of the lexico-grammatical categories of *hedges* was more by the left-wing compared to the right-wing. The right-wing commenters never used a single noun, whereas the left-wing commenters used 11.67 per 10,000 words. The left-wing commenters used verbs, adjectives, adverbs, and phrases more than the right-wing, while the right-wing commenters used modal verbs more in comparison to the left-wing. While using the adverbs and phrases, it was noticed that the right-wing commenters used adverbs and phrases less frequently, whereas the left-wing commenters used them more (almost 2 times and 3 times), respectively. The usage difference among the left-wing and right-wing sub-corpora indicates that the left-wing commenters of the UK sub-corpora were more eager not to take complete responsibility; preferably, they tried to avoid the situations politely. On the contrary, the right-wing commenters used modal verbs which suggests they were more eager to propose a solution to the problem or give a kind of suggestion rather than completely refraining from taking responsibility.

The USA: The quantitative analysis of *hedges* used by the USA commenters reveals that overall the commenters of the sports domain used the lexico-grammatical categories more, followed by the entertainment and politics domain. In the sports domain, the overall usage of the lexico-grammatical categories of *hedges* among the left-wing and right-wing sub-corpora was almost similar, with a slight difference. On the contrary, in the politics and entertainment domains, the right-wing commenters used the

lexico-grammatical categories of *hedges* more than the left-wing. We noted that all the lexico-grammatical categories of *hedges* were used by the commenters, among which modal verbs were used most frequently, and nouns and phrases were used less frequently in all three domains. However, a notable difference was observed in the overall usage of the lexico-grammatical categories of *hedges* among the two political ideologies of the entertainment and politics domains. Unlike the sports domain, in the entertainment and politics domains, nouns, verbs, adverbs, and phrases were observed to be used more by the right-wing commenters than the left-wing. It was noticed that both the left-wing and right-wing commenters of the entertainment domains used modal verbs almost the same, whereas the right-wing commenters of the politics domain used them significantly more (98.11 per 10,000 words) than the left-wing (57.93 per 10,000 words). Although the use of hedges-phrases was much less among the three domains, we observed that the commenters of the sports domain used them more (24.33 per 10,000 words), followed by the politics and entertainment domains (16.46 and 14.46 per 10,000 words, respectively). The frequent use of the lexico-grammatical categories of *hedges* points out the fact that the USA commenters, irrespective of domains and political ideology, were able to withhold the proposition strongly in front of the readers and commenters.

Occurrences in Lexico-Grammatical categories

Furthermore, in Table 7.2, we provide more specific details of the lexico-grammatical categories of *hedges*, more specifically, we picked up the five most frequent words of each lexico-grammatical category (verb, noun, adverb, adjective, phrase, and modal verb). Here, we mention the most frequent words that appeared in the UK, India, and the USA sub-corpora and discuss with examples along with the sense in which it is used in the comments by the commenters.

Verbs: Among the verbs, ‘want’, and ‘think’ are the most frequently used verbs by the commenters of the three nations. While the India commenters used the lexical verb

Table 7.2: Top-5 (Normalised frequency per 10,000 words) words of lexico-grammatical categories of *hedges*

	INDIA		UK		USA	
Verb	want	107.72	think	152.27	think	153.47
	think	103.69	want	145.39	want	87.17
	feel	51.85	see	22.97	feel	24.78
	see	12.76	feel	21.12	assume	8.50
	expect	9.06	expect	10.29	attempt	7.85
Noun	tendency	3.85	idea	29.58	idea	26.08
	idea	2.23	view	13.90	view	13.40
	view	2.16	argument	8.65	argument	12.33
	belief	2.16	perspective	6.98	expectation	3.75
	expectation	1.81	possibility	4.15	belief	3.11
Adverb	just	150.40	just	204.04	just	165.19
	maybe	10.75	probably	54.35	maybe	38.10
	sometime	9.67	maybe	45.73	probably	35.33
	quite	7.22	quite	30.99	quite	15.68
	almost	5.53	perhaps	26.59	often	14.27
Adjectives	some	91.20	some	92.46	any	111.77
	any	77.84	any	66.52	some	108.36
	anyone	16.44	anyone	18.57	likely	17.03
	possible	8.20	common	7.77	anyone	15.81
	common	5.54	likely	5.63	certain	12.34
Phrases	like	19.01	like	49.47	at least	25.43
	at least	7.44	at least	17.91	like	22.86
	in general	2.77	in general	2.60	in general	5.88
			in theory	1.67	in theory	1.08
			in part	1.67		
Model Verb	can	215.13	can	175.91	can	193.88
	may	85.20	may	63.25	may	68.04
	shall	157.01	shall	98.71	shall	94.35
	will	98.82	will	288.61	will	190.03

‘think’ for 103.69 per 10,000 words, the commenters of the UK and USA used it almost similarly. Again, in the use of another lexical verb ‘want’, we noticed that the UK commenters used it more frequently, followed by the India and USA commenters. These verbs also act as epistemic verbs, which project the writer’s faith about a proposition and also express the writer’s reluctance to assure something. In the case of the verb ‘feel’, the India commenters were seen to use it more frequently, followed by the USA and UK commenters. The other two verbs, ‘see’ and ‘expect’, were found in the India and UK sub-corpora; however, the verbs ‘assume’ and ‘attempt’ were found in the American sub-corpora. Now, we mention some examples of the lexical verb ‘want’ from the UK, India and USA sub-corpora.

Example 7.5.1 *“The same Pep that just got jobbed by Liverpool? The same Pep that*

has never won the Champions League without Messi? That Pep? He best buy Messi, Ronny and Neymar if he wants to win a champions league with the small dirty oil club!!” (source:UK-RW-SPR)

In this comment from the UK sub-corpora, the commenter spoke about the former Spanish footballer and football manager, Pep Guardiola. According to the commenter, if Pep Guardiola wants to win the Champions League, he must buy Messi, Ronny, and Neymar (3 famous footballers). Here, the commenter used the lexical verb ‘want’ to mean something desired. It is more of an advice from the commenter’s side rather than a strict assertion. Also, it can be referred to as a solution that the commenter was prescribing for the welfare of the club.

Example 7.5.2 *“The Hindu.. Heading could have been India wins Gold and Silver in CWG. The article gives a feeling of bias to Saina. It’s true that Saina came back strong, but isn’t it important for a newspaper like The Hindu, to stand out from others. Both are equally important to us and should be promoted equally. It shouldn’t be like other media’s who want to see a fight between them. There is only one winner in a match. But for India it was two in this match. Great going Saina and Sindhu. Keep winning! Nation is proud!” (source:IND-LW-SPR)*

In this example from the India sub-corpora, the commenter talked about two famous Indian badminton players named Saina Nehwal and P.V Sindhu. While Saina Nehwal won gold in the Commonwealth Games for India, P.V Sindhu won silver. The commenter here argued over the role of India’s popular newspaper named *The Hindu* to represent both the players equally with the same respect; however, according to the commenter, *The Hindu* highlighted Saina Nehwal more than P.V Sindhu. Further, he criticized the role of other newspapers also which were more eager to see a brawl between the players. So, it can be said that the verb ‘want’ here refers to something that is desired. The commenter talks about probability, although he is not certain about the idea. The verb ‘want’ was used by the commenter to express a kind of vagueness and uncertainty from the commenter’s side.

Example 7.5.3 “*VERY hypocritical guards... Steven Spielberg is anti-gun, but has possibly the largest private gun collection in the US. including machine guns, I suspect. But has he surrendered them to the government so they can be destroyed? No. But he wants everyone else to do that.*” (*source:USA-RW-ENT*)

In this example from the USA sub-corpora, the commenter criticized the famous American director and producer Steven Spielberg who, despite having a large number of private gun collections and being anti-gun, has not surrendered his guns to the government. The commenter was even more surprised that the actor advised others to give up on their guns even after possessing guns himself. Here, the lexical verb ‘want’ refers to something Steven Spielberg desires for. The commenter upholds the actor’s voice, suggesting the common people give up on their guns. Although the commenter here is stating something that is to be done, he withholds his full responsibility to the idea.

Nouns: Among the nouns, the commenters of the UK, India, and the USA used ‘idea’ and ‘view’ most frequently, although there was a difference in the frequency. Whereas the India commenters used the noun ‘idea’ less frequently (2.23 per 10,000 words), it was used significantly by the UK and the USA commenters. The other nouns mostly used in all the three sub-corpora are *tendency, belief, expectation, argument, perspective, and possibility* (See Table 7.2). Here, we cite some examples of the comments containing the noun ‘idea’ from the UK, India, and the USA sub-corpora:

Example 7.5.4 “*France, Germany and Poland have signed some huge trade deals recently with Commonwealth countries such as India. The idea that the UK cannot do trade deals inside the EU is false.*” (*source:UK-RW-POL:2*)

In Example 7.5.4 from the UK sub-corpora, the commenter used the noun *idea* to refer to an understanding or thought. The commenter described an incident where the Western countries like France, Germany, and Poland had signed big trade deals with the commonwealth nation India. So, according to the commenter UK can also perform

trade deals inside the EU even after Brexit. Here the noun ‘idea’ was used in a positive sense. It is more of a fact than a suggestion by the commenter’s side.

Example 7.5.5 *“Afridi u r nevr cald in IPL in last 9 sessions. Bogus u r kicked out from IPL. and y we will call u to play, neither u bowl bat or field well. U perform in one out of 100 match. U r d most inconcistent player world has ever seen. Ur pakistan team should use u as a water and towel carrier in between matches, actually a good idea. we can cal u for a spot boy in IPL but not this sessions as already for spot also there is huge rush in IPL. So better luck next time and try practicing hard for a spot boy post next session.” (source:IND-LW-SPR)*

In this example from the Indian sub-corpora, the commenter criticized Shahid Afridi (a Pakistani cricketer) for being out of form, and that was the reason for not calling him into the IPL (Indian Premier League), which is a popular cricket tournament. While criticizing Shahid Afridi the commenter used the noun ‘idea’, which has a negative connotation. It is used to refer to the commenter’s subjective opinion. However, despite showing subjectivity, the noun ‘idea’ is not showing the commenter’s full certainty about the proposition.

Example 7.5.6 *“I would advise the president to consider taking care of the immigration issues as he campaigned before jumping into the complex morass of capital punishment for drug smugglers. After all, as I understood his previous positions on drug use, he is all for legalizing marijuana. I thought this was position on drugs in general - legalize and tax. Given the consequences of alcohol. I think the idea a dubious solution. While I am all for tough renders on drugging the country any further, we might want to limit access to our border via tunnels under the nonexistent wall.” (source:USA-RW-POL:1)*

In this example from the USA sub-corpora, the commenter used the noun ‘idea’ to refer to an understanding or thought. The commenter talked about the actions

of Donald Trump (the then President of the USA). According to the commenter, the President must focus on handling the immigration issue and then decide on capital punishment for drug smugglers. He also stated that legalizing alcohol is an uncertain idea compared to legalizing drugs in the USA. Here, the noun ‘idea’ was used negatively as the commenter expressed doubt or hesitancy about the notion. It expresses uncertainty from the commenter’s point of view.

Adverbs: Table 7.2 depicts that among the adverbs, the most frequently used adverb by the UK, India, and USA commenters was ‘just’. The UK commenters used it most frequently, followed by the USA and India commenters. Another two adverbs, ‘may be’ and ‘quite,’ were also seen in all three sub-corpora. Both these adverbs were mostly used by the UK commenters, followed by the USA and India commenters. The other adverbs used by the commenters of all three nations are *sometime, almost, probably, perhaps, and often*. Here we show some examples of the adverb ‘just’ from our corpus.

Example 7.5.7 “*Ah, the internet is just a passing fad, probably won’t be around this time next week. Black and white silent movies are the wave of the future.*” (**source:UK-LW-ENT:1**)

In this example from the UK sub-corpora, the commenter argued about the future of the OTT (Over the Top) platforms where movie content can be seen mainly by using an internet connection or cable. According to the commenter, the internet is only a craze that will eventually pass, and in the future, black and white silent movies will reign. The adverb ‘just’ was used to reduce the statement’s force to suggest something that it is not very important. By using ‘just’, the commenter was emphasizing subjectivity. It was a kind of opinion from the commenter’s side that was eventually reducing the force of the comment.

Example 7.5.8 “*...The people Killed Indira Gandhi were not Sikhs but Terrorists. Those who kill and bombs do not belong to any religion, they just misuse religion.*”

Terrorists are cowards. Extremist do not follow any religion. They interfere in other people religion to create hatred and blood bath.” (source:IND-LW-POL)

In this comment from the India sub-corpora, the commenter refers to the gruesome act of the assassination of Indira Gandhi (the former Prime Minister of India). He referred to this incident to claim his thoughts which emphasize that extremism does not follow any religion and that extremists use religion as a tool to accomplish evil acts. Here, the adverb ‘just’ was used to reduce the statement’s force and decrease its seriousness. It was used in a negative sense by the commenter.

Example 7.5.9 *“Also, did my own experience growing up in a town with terrible, awful, no good cops in the 80s was that everyone would join in to condemn the cops when the object of their ill intentions was white but would turn right around and say “They’ re just doing their jobs” when the object was black.” (source:USA-RW-POL:2)*

The commenter of the USA sub-corpora distinguished the role played by the police officers in the 80s and the role played by the people of the commenter’s neighbourhood. According to him, the attitude of the common people of his neighbourhood changed towards the policemen when they had to deal with a black person. Here, the adverb ‘just’ was used to reduce the force of the commenter’s statement. It refers to something casual that eventually decreases the severity of the incident.

Adjectives: While calculating the adjectives, we found that three adjectives, ‘some’, ‘any’, and ‘anyone’, were used by the India, UK, and USA commenters. The USA commenters used ‘some’ and ‘any’ more significantly in their comments. On the other hand, ‘some’ was used more by the UK commenters in comparison to the India commenters, whereas the adjective ‘any’ was used more by the Indians than the UK commenters. Among the other adjectives, *common*, *likely*, *certain* were used by the commenters of all three nations. In order to grab the readers’ attention, the modal adjectives generate some power for the nouns. Here, we mention some examples of the adjective ‘some’ from the UK, India, and the USA sub-corpora.

Example 7.5.10 *“Morrison wants a Hard Brexit. Maximized job losses, maximized falls in output, in growth, in welfare, and in government revenues. Have some compassion for your fellow countrymen and women, and especially for those less fortunate yourself, man; and stop the fake displays of ignorance and stupidity.” (source:UK-RW-POL:3)*

Example 7.5.11 *“Rohit Sharma is failing in all formats nowadays. He should take a rest for some time. Congratulations SRH for winning 2 games out of 2.” (source:IND-RW-SPR)*

Example 7.5.12 *“While I agree with you that these fine students are motivated by their tragedies, I am sure that there are some people that are trying to groom them, trying to make sure that they do not fall into the traps set by those that are against them I’m ok with that grooming because these students need to make sure that they are reminded of what they’re fighting.” (source:USA-LW-ENT)*

In example 7.5.10, the UK commenter criticized the UK politician Morrison who wanted Brexit, and as a result, the people of the UK had seen a downfall in every sector. Further, the commenter criticized the UK politician and asked him to have some sympathy for his fellow citizens. Here, the quantitative adjective ‘some’ was used to mean an unspecified amount of emotion, as emotion cannot be counted. The commenter used the adjective in a suggestive manner. It was more of a suggestion than an assertion. In example 7.5.11 from the India sub-corpora, the commenter criticized the Indian cricket player Rohit Sharma, who, according to the commenter, was out of form at that time, and so he must rest for a while. The quantitative adjective ‘some’ was used to mean an unspecified period of time. It has a negative connotation that upholds subjectivity on the commenter’s part. The commenter from the USA sub-corpora in example 7.5.12 explained the tragic situation where some students took weapons in their hands due to the influence of the surroundings. The commenter was hopeful that there are still a certain amount of good people in society who would help

the unfortunate students return to the main track of their lives. Again, the quantitative adjective ‘some’ was used here to mean an unspecified amount of people.

Phrases: In the use of phrases, it was noticed that ‘like’, ‘at least’, and ‘in general’ were most frequently used by the UK, India, and USA commenters. The UK commenters used ‘like’ comparatively more, followed by the USA and India commenters. The USA commenters used ‘at least’ more, followed by the UK and India commenters. Although ‘in general’ was used considerably by the USA commenters, it was used almost the same by the India and UK commenters. It was observed that in the India sub-corpora, only these three adjectives were used; however, in the USA and the UK sub-corpora, two other phrases, ‘in theory’ and ‘in part’, were used, respectively. Here, we present some examples from the UK, India, and the USA sub-corpora.

Example 7.5.13 *“Queer is an accepted term for non normative sexualities and is used academically and has been for about 40 years at least.” (source:UK-LW-ENT:2)*

Example 7.5.14 *“Shock to hear of death, but very strange–no illness... At least Dubai police will get to the bottom of this death, as many actress die in India and no one checks properly on how they died.. and a cremated very quickly.... NO COVER UP THIS TIME.” (source:IND-LW-ENT)*

Example 7.5.15 *“If nothing else, Carter did at least one great thing. He looked America in the eye and warned us about depending on the Middle East for energy. It’s not on him that no one listened or cared.” (source:USA-LW-POL)*

In example 7.5.13, the commenter described that the term queer (people who are not heterosexual) is accepted and used academically for a minimum of 40 years. The commenter used the phrase ‘at least’ to emphasize that something is good in a bad situation. Here, the commenter expressed his personal thoughts while using the phrase ‘at least’. The use of this phrase also added a sense of subjectivity on the part of the commenter, who tried to guide the readers (in our case, the other commenters).

In example 7.5.14, the commenter discussed the cause of Sridevi's (a famous Indian film actress) death, and he was hopeful that as Sridevi died in Dubai, at least the cause of her death would be properly investigated by Dubai police, unlike most of the death mysteries of the actors that went unresolved in India. Here, the phrase 'at least' emphasizes that something is good in a bad situation. Here, the commenter was describing something that he was hopeful for. It also expressed a kind of vagueness as the commenter himself was not entirely certain, although he is hopeful for that. Also, in example 7.5.15, the commenter praised the foresight of the President (39th) of the United States, Jimmy Carter, who warned about the dependency of the United States on the Middle East for energy. He used the phrase 'at least' to emphasize something positive in a negative situation. The commenter expressed his personal belief by using the phrase 'at least'.

Modal verbs: In our corpus, four modal verbs (*can*, *may*, *shall* and *will*) were used by the UK, India, and the USA commenters that acted as hedging devices. The writers used these devices to project their hesitation and uncertainty about a particular proposition. Also, writers used modal verbs such as 'can' and 'may' to express possibility. In previous studies, such as Carrió-Pastor (2019b), Siddique et al. (2018), modal verbs were frequently used by the writers. The modal verb 'can' was used enormously by the commenters of all three nations. It was most frequently used by the India commenters, followed by the USA and UK commenters. The modal verb 'may' and 'shall' were used more by the India commenters; however, they were used almost the same by the UK and USA ones. On the contrary, 'will' was used significantly less by the India commenters, whereas it was used enormously by the UK and USA commenters. The following examples demonstrate the use of hedges modal verbs from the UK, India, and the USA sub-corpora.

Example 7.5.16 *“On Saturday, June 23rd (the second anniversary of the Brexit vote day), at ten am... if you still believe in Brexit.. simply stand outside your door.. where you can be seen, for ten minutes... to demonstrate that we have NOT changed*

our minds... and that we are, in fact, not afraid to LEAVE....And if you originally voted to remain, but have since' seen the light take up this campaign and help spread the word.. but, in case it doesn't, PLEASE spread the word yourself on whatever media you support. we'll call it "not afraid to leave" day. Maybe someone who understands twitter can start a hashtag thing..and please copy and paste wherever you can." (source:UK-RW-POL:1)

Example 7.5.17 *"Great! All india finals! Wish them and others to reap laurels for our nation. The potential is there. If properly utilised, our players and athletes can achieve in the Olympics also." (source:IND-LW-SPR)*

Example 7.5.18 *"This is an example of what is wrong with most conservative thinking. They don't understand peoples' motives when they have NOTHING to do with money and power. Trump was elected by promising that he could make struggling people rich, "ONLY I can save you. "When Trump and his surrogates know full well they couldn't care Less about those people, who will end up even worse off than when they started just like in the W.Bush recession. While Trump and friends walk away with a tax dollars. Their ONLY motivation is money and power." (source:USA-LW-ENT)*

In example 7.5.16, the UK commenter asked his fellow commenters to act to support Brexit unitedly. In this comment, the modal verb 'can' was used three times by the commenter to express possibility. In all three cases, the commenter used 'can' to suggest something to his fellow commenters. It was more of an opinion from the commenter's side rather than a certain assertion. Again, in example 7.5.17, the Indian commenter talked about the excellent performance of the Indian players in the Commonwealth Games, and according to him, the Indian players have great potential to win even in the Olympics if adequately utilized. Here, the modal verb 'can' was used to express possibility. Also, it emphasized the commenter's personal beliefs and helped him to express his thoughts in an indirect way. In example 7.5.18, the USA commenter used the modal verb 'can' to express possibility. Here, the commenter referred to the

then President of the USA, Donald Trump, who, according to the commenter, was elected because he promised to make the lives better of the struggling people. In this comment, the modal verb ‘can’ indicates possibility. However, it reflects hesitancy on the commenter’s part, who was not entirely sure about the assertion.

Comparisons of the occurrences of *hedges*

From Figure 7.7, we extracted some notable characteristics of using *hedges* in the comments by the commenters of India, the UK, and the USA.

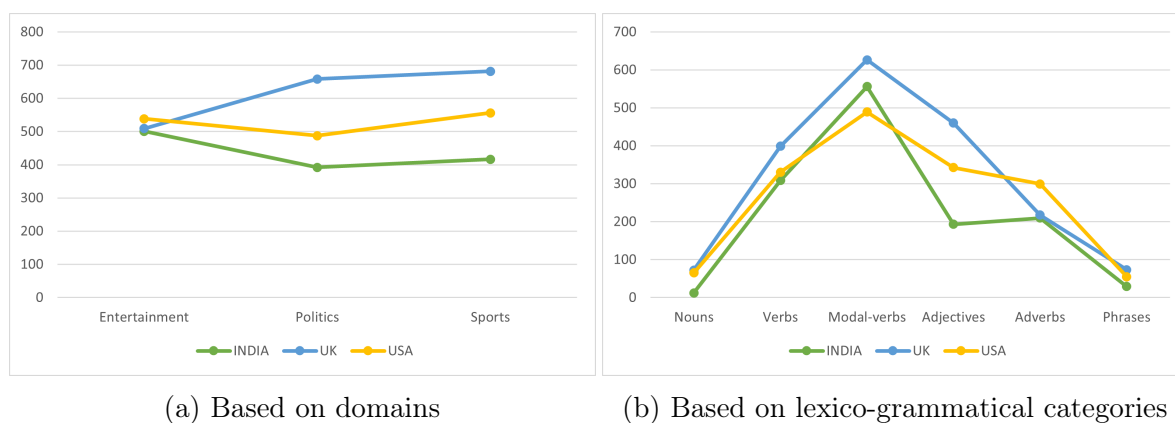


Figure 7.7: Comparisons of the occurrences (normalized) of *hedges*

- From Figure 7.7a, it can be seen that the UK commenters employed *hedges* in politics and sports domains significantly more than the commenters of the USA and India, which indicates that the UK commenters were not so confident enough to say something certain about a proposition. Rather they were interested more in avoiding full commitment and responsibility towards an idea. Interestingly, the India and USA commenters were more vocal in projecting their subjective opinions about a particular idea or proposition in the entertainment and sports domains, compared to the politics domain. The UK and the USA commenters employed *hedges* in the sports domain the most; however, their usage was significantly different. While commenting on entertainment-related topics, the usage of *hedges* was seen to be most frequent in the India sub-corpora, followed by the

USA and UK which suggests that the India commenters of the entertainment domain tended to show politeness by not asserting something certainty, rather they were more eager to provide alternative solutions and suggestions.

- Analysing our data based on the lexico-grammatical categories of *hedges*, we conclude from the Figure 7.7b that modal verbs were employed most frequently by all the commenters irrespective of their region which indicates that these modal verbs helped the commenters to express their thoughts and ideas indirectly in front of the other commenters. Notably, the usage pattern of the lexico-grammatical categories of *hedges* was almost the same. Hedges-nouns and phrases were used less frequently by all the commenters; more specifically, the India commenters were reluctant to use hedges-nouns in their comments, irrespective of domain. In the use of hedges-adverbs, the USA commenters used them well ahead of the UK and India commenters, whereas the UK and India commenters used them almost the same. Another likeness between the commenters of the USA and India was the use of hedges-verb and hedges-phrases in their comments.

7.5.2 Interactional: Boosters

Boosters are one of the interactional metadiscourse markers used by writers to strengthen the commitments of writers to propositions. Unlike *hedges*, *boosters* refer to the communicative strategies used by the writers to express commitment to propositions to persuade the readers of their authenticity. These markers help the writers to express their firm commitments and to express certainty and validity to propositions (Hyland (2005b), Cao and Hu (2014), Peacock (2006). Previous studies Hyland (1998a), Peacock (2006), Skorczynska and Carrió-Pastor (2021), Yazdani et al. (2014) have investigated the use of *boosters*. In our study, *boosters* played the following roles: 1) in expressing the commenter's commitment and certainty to a specific proposition or argument, 2) to show unanimity with the other commenters by stressing the truth, 3) to show the writer's confidence regarding a particular proposition or fact by increasing the tone of

the claims. Table 7.3 depicts the normalized frequency of *boosters* found in the UK, India, and USA sub-corpora. This section describes the in-depth analysis of *boosters* from our corpus. Similar to the studies of Hyland (2005b) and Mur-Dueñas (2011), where *boosters* were used comparatively less than *hedges*, in our corpus also, *boosters* were used comparatively less than *hedges* by the commenters. We demonstrated the frequency pattern of using *boosters* from our corpus according to the domains and political ideology. There are five lexico-grammatical categories of *boosters*: nouns, verbs, adjectives, adverbs, and phrases. Here, we also cite some examples of the occurrences from the UK, India, and the USA sub-corpora.

Table 7.3: Normalized frequencies of *boosters* found in India, the UK, and the USA

	Entertainment			Politics			Sports			
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall	
INDIA	Nouns	0.00	5.31	5.31	7.57	7.25	14.82	0.00	0.00	0.00
	Verbs	22.17	26.54	48.71	37.85	38.05	75.90	15.58	14.02	29.60
	Adjectives	0.00	0.00	0.00	3.24	1.81	5.06	0.00	9.35	9.35
	Adverbs	33.26	15.92	49.18	16.22	10.87	27.09	15.58	37.38	52.96
	Phrases	0.00	0.00	0.00	4.33	3.62	7.95	2.23	0.00	2.23
	Total	55.43	47.77	103.20	69.21	61.61	130.82	33.39	60.75	94.13
UK	Nouns	0.00	7.70	7.70	23.32	17.31	40.63	4.98	10.61	15.59
	Verbs	26.67	19.25	45.91	30.31	31.35	61.66	34.87	31.84	66.71
	Adjectives	3.33	0.00	3.33	11.66	6.08	17.74	2.99	2.65	5.64
	Adverbs	23.33	23.09	46.43	18.65	26.67	45.32	24.91	15.92	40.83
	Phrases	3.33	3.85	7.18	2.33	6.08	8.41	4.98	15.92	20.90
	Total	56.67	53.89	110.55	86.28	87.49	173.77	72.73	76.94	149.67
USA	Nouns	12.46	8.23	20.69	18.74	8.47	27.22	11.86	4.31	16.17
	Verbs	24.15	28.79	52.94	28.40	28.99	57.39	37.27	30.17	67.44
	Adjectives	1.56	2.06	3.61	0.57	4.46	5.03	8.47	2.16	10.63
	Adverbs	20.25	30.85	51.10	21.01	40.14	61.15	18.63	29.09	47.73
	Phrases	3.12	4.11	7.23	2.84	4.46	7.30	3.39	10.78	14.16
	Total	61.53	74.04	135.57	71.56	86.52	158.08	79.62	76.51	156.13

The India, UK, and USA commenters of the three domains (entertainment, politics, and sports) used the following *boosters*: nouns (*fact, evidence, majority, conclusion, assertions*), verbs (*know, show, demonstrate, establish, hold, confirm, prove, conclude, stress, determine, reveal, highlight, assert*), adverbs (*always, actually, particularly, clearly, especially, indeed, highly, strongly, substantially, essentially, generally, mostly, largely, fully, entirely, widely, consistently, primarily, constantly*), adjectives (*clear, vast, extraordinary, evident, substantial, thrilling*) and phrases (*of course, in fact, for the most part, in effect*).

India: With respect to different domains, it is evident from Table 7.3 that the overall usage of the lexico-grammatical categories of *boosters* was frequently seen more in the politics domain of the India sub-corpora followed by the entertainment and sports domains. In the politics domain, overall, the verbs were used more, and the adjectives were used less frequently by the commenters. In the use of nouns and verbs, both the left-wing and right-wing commenters used nouns almost in the same manner. Regarding the use of adjectives, adverbs, and phrases, the left-wing commenters used more compared to the right-wing ones. So, if we observe the overall frequency of using the lexico-grammatical categories of *boosters*, we can see that the left-wing commenters used them a bit more than the right-wing commenters which suggests that the left-wing commenters of the politics domain were more eager to emphasize certainty regarding an idea or thought. Interestingly, the India commenters of the entertainment domain used no adjectives and phrases in their comments. A notable difference was noticed in the use of nouns and adverbs. While there was not a single use of noun by the left-wing commenters of the India sub-corpora, the right-wing commenters used it for 5.31 per 10,000 words. Again, in the use of adverbs, the right-wing commenters used it for 15.92 per 10,000 words, and the left-wing commenters used it for 33.26 per 10,000 words, almost more than 2 times. In the overall use of the lexico-grammatical categories of *boosters*, the left-wing commenters were seen to use them more than the right-wing. In the sports domain, overall verbs and adverbs were used most by both the left-wing and right-wing commenters. However, not a single use of the noun was noticed by the left-wing and right-wing commenters; the verbs were seen to be used almost the same by both the left-wing and right-wing. Also, no adjective was used by the left-wing commenters, whereas the right-wing commenters of this domain used adjectives for 9.35 per 10,000 words. In the use of adverbs, the right-wing commenters used adverbs more than two times than the left-wing. A slight amount of phrases were used by the left-wing, whereas the right-wing commenters never used them a single time.

The UK: Concerning different domains, the lexico-grammatical categories of *boost-*

ers were used mostly in the politics domain, followed by the sports and entertainment domain. In the politics domain, verbs and adverbs were used most frequently, and phrases were less frequently used by both the left-wing and right-wing commenters. Regarding the use of verbs, adverbs, and phrases, the right-wing commenters used them more in comparison to the left-wing ones. On the contrary, in the use of nouns and adjectives, the left-wing commenters used them almost 2 times more than the right-wing. In the use of phrases, the left-wing commenters used it for 2.33 per 10,000 words, whereas the right-wing commenters used it for 6.08, which is almost three times more. In the sports domain also, verbs and adverbs were mostly used irrespective of political ideology. In the use of nouns and phrases, we can see from Table 7.3 that the right-wing commenters used nouns and phrases almost more than twice than the left-wing. However, in the use of verbs and adjectives, both the left-wing and right-wing commenters used them almost the same. Regarding the use of adverbs, the left-wing commenters used more than the right-wing ones. In the entertainment domain, the commenters most frequently used verbs and adverbs in the entertainment domain, irrespective of political ideology. The left-wing commenters never used a single noun, whereas the right-wing ones used them for 7.70 per 10,000 words. In the use of verbs, adjectives, and adverbs, it was noticed that the left-wing used comparatively more than the right-wing. Both the left-wing and right-wing commenters used phrases for almost the same.

The USA: Among the three domains, the commenters of the politics and sports domain overall used the lexico-grammatical categories of boosters almost the same. Among all three domains, verbs, and adverbs were used more compared to the other lexico-grammatical categories of boosters. In the politics domain, nouns were used most frequently by the left-wing compared to the right-wing, which was almost 2 times more. Irrespective of political ideology, verbs were used almost the same. In the use of adjectives, the use was significantly less frequent for both the left-wing and right-wing, more specifically in the left-wing (0.57 per 10,000 words) and the right-wing

(4.46 per 10,000 words). In the use of adverbs and phrases, the right-wing commenters were seen to use them almost twice more compared to the left-wing. Concerning the sports domain, while the left-wing commenters used nouns, verbs, and adjectives more than the right-wing, adverbs and phrases were used by the right-wing commenters more than the left-wing ones. Adjectives were used by the left-wing commenters 4 times more than the right-wing. A notable difference was noticed when the left-wing used phrases much less (3.39 per 10,000 words), and the right-wing commenters used them almost 3 times more (10.78 per 10,000 words). The quantitative analysis of the entertainment domain shows that except nouns, all the other lexico-grammatical categories of *boosters* were used more frequently by the right-wing commenters than the left-wing ones that indicates the right-wing commenters' were confident enough to express their certainty about a proposition. On the other hand, the left-wing commenters used nouns more than the right-wing.

Occurrences in Lexico-Grammatical categories

To explore the lexico-grammatical categories further, we delve deep into the quantitative analysis of the lexico-grammatical categories of *boosters* in detail. Table 7.4 depicts the normalized frequency of the top five occurrences from each of the lexico-grammatical categories. Also, we provide examples that help to understand the use of the lexico-grammatical categories of *boosters* by the commenters.

Verbs: The common booster-verbs that occurred in all the three sub-corpora (India, UK, and USA) were 'know', 'show', and 'establish'. The verb 'know' was used mostly by the UK commenters, followed by the USA and India commenters. The India commenters used the verb 'show' more than the USA and the UK commenters. The verb 'establish' was used almost the same in all three sub-corpora. These three verbs acts as boosting devices to convey certainty from the part of the commenter. The other most frequently used verbs were *prove*, *highlight*, *hold*, *confirm*, and *demonstrate*. However, we noticed a difference in the use of these verbs by the commenters in the

three sub-corpora. Here, we demonstrate examples from the UK, India, and the USA sub-corpora.

Table 7.4: Top-5 Normalised frequency per 10,000 words of lexico-grammatical categories of *boosters*

	INDIA		UK		USA	
Verb	know	64.48	know	122.95	know	111.44
	show	56.06	show	22.48	show	35.81
	prove	13.48	hold	9.88	prove	10.65
	highlight	5.00	confirm	7.97	demonstrate	4.76
	establish	3.85	establish	3.60	establish	3.81
Noun	evidence	7.25	fact	33.44	fact	44.73
	fact	6.98	majority	15.47	majority	11.52
	majority	5.90	evidence	11.97	evidence	6.29
			conclusion	2.57	assertions	1.08
			assertions	0.47	conclusion	0.45
Adverb	always	84.43	actually	41.84	actually	50.05
	actually	9.79	always	32.02	always	46.64
	especially	7.57	clearly	11.82	especially	12.32
	highly	7.44	especially	10.51	clearly	9.31
	particularly	5.75	particularly	10.09	mostly	5.97
Adjectives	thrilling	9.35	clear	20.35	clear	13.73
	clear	2.16	vast	3.27	substantial	2.06
	substantial	1.81	evident	1.93	extraordinary	1.69
	vast	1.08	extraordinary	1.17	vast	1.34
			substantial	0.00	evident	0.45
Phrases	Of course	8.01	Of course	30.72	Of course	16.21
	In fact	2.16	In fact	5.78	In fact	7.55
					For the most part	3.40
					In effect	1.52

Example 7.5.19 *“Donald Trump is a passing phase in American Politics. Everyone knows the President has no real power (for example Obama could get zilch legislation through even on gun controls). The relationship between America and the UK however has stood the test of time - we can always rely on the USA and we need them as our best buddies.” (source:UK-LW-POL:2)*

In this example from the UK sub-corpora, the commenter used the booster verb ‘know’ while criticizing the power of the then President of America, Donald Trump. Here, the verb ‘know’ means being aware of something through observation. According to the commenter, everyone is aware that the President has no real power, and the verb ‘know’ express certainty from the commenter’s part. It has a convincing tone which is

used by the commenter to persuade the other commenters. Also, it conveys a kind of confidence from the author about his idea.

Example 7.5.20 *“Since only Bollywood, Cricket and Politics are being fed to Indians post Independence, only such deaths create shockwaves here. The high fly club who the common Indian rarely see is off Twitter and Instagram. Know how many people living in the same city die like dropping apples everyday on trains...” (source:IND-LW-ENT)*

In this example from the India sub-corpora, the commenter used the booster verb ‘know’ while questioning his fellow commenter. The commenter referred to the death of the famous Indian film actress Sridevi which entirely shook the nation. Here, the commenter criticized the fact that in spite of projecting the hardships of the common people, the news from the film industry, politics, and cricket are shown. So, the death news of such a superstar creates disorder in the country. The verb ‘know’ emphasizes the endeavour from the commenter’s side to convince the readers (in our case, the other commenters) about the idea that he is confident about.

Example 7.5.21 *“Do we even Know if Trump has a diploma? I believe he always say, “I attended...“Never, “I graduated from...” ” (source:USA-LW-POL)*

In this comment from the USA sub-corpora, the commenter asked a question regarding the educational qualification of the then President of America, Donald Trump. In this regard, he used the verb ‘know’ to ask if the citizens of America have the necessary knowledge of Trump’s educational qualifications. Here, the commenter wanted to project the implicit truth of Trump’s educational qualification and in order to stress his claim, he showed solidarity by stressing the truth. Therefore, it can be stated that the verb ‘know’ serves as a boosting device by the commenter to exhibit solidarity with the readers (in our case, the other commenters).

Nouns: Three nouns were common in the India, UK, and USA sub-corpora: ‘evidence’, ‘fact’, and ‘majority’. All these nouns were used less by the India commenters

compared to the UK and USA. Another two nouns, ‘conclusion’ and ‘assertion’ were seen to be used by the UK and the USA commenters. The USA commenters used the noun ‘conclusion’ less frequently; however, the UK commenters used them a bit more. Contrarily, the noun ‘assertion’ was used much less by the UK commenters, whereas it was used a bit more by the USA ones. Here, we present examples of using the booster noun ‘fact’ from the three sub-corpora (UK, India, and USA).

Example 7.5.22 *“The visa costs are for ALL countries outside of Schengen so try getting your facts right -even the Irish will need visas (no doubt they will try and get the UK to pay for them!!)” (source:UK-RW-POL:2)*

Example 7.5.23 *“Well we in India will always celebrate history and archaeological facts and accept scientific truth. Muslims have been in the game of erasing history for hundreds of years. If one go to any arab country, muslims have destroyed all pre islamic archaeology, art, books and history. For all they know is that history started with Muhammad...” (source:IND-LW-POL)*

Example 7.5.24 *“The most appalling aspect of Donald Trump’s speeches and tweets this weekend is the fact that over 3 million Americans in Puerto Rico are truly suffering in the wake of Hurricane Maria and are in dire need of assistance. Instead of tweeting at Steph Curry and complaining about football players kneeling for the national anthem, why not tweet or speak about things your administration...” (source:USA-RW-POL:1)*

In the examples mentioned above from the three sub-corpora (UK, India, and USA), the noun ‘fact’ was used by the respective commenters to mean something that is known to have happened or to exist, especially something for which proof exists, or about which there is information. In the first comment from the UK sub-corpora, the commenter talked about a situation after Brexit and asked his fellow commenters to know the truth first before commenting on something. Here, the noun ‘fact’ was used

as a booster to emphasize something the commenter was certain about. In the second example from the India sub-corpora, the commenter described that Indians believe the reality or truth that is known or proved to be true. Likewise the previous example, the commenter here tried to emphasize something that is known to exist. In the third example from the USA sub-corpora, the commenter used the noun ‘fact’ to summarize Donald Trump’s speeches and tweets, whereas the commenter used this noun to refer to something true or real. Also, it refers to something that already exists and is known.

Adverbs: In all three sub-corpora, the three common adverbs ‘always’, ‘actually’, and ‘especially’ were used more frequently. The India commenters used ‘always’ most, followed by the USA and UK commenters. In the use of ‘actually,’ we noticed that the India commenters used them very less, whereas it was used more frequently by the UK and USA commenters (almost 4 times more). The other highly used adverbs were *highly, particularly, clearly, and mostly*. Here are some examples of the adverb ‘always’ from the UK, India, and USA sub-corpora.

Example 7.5.25 *“I have always admired Christopher Booker and on opening the papers on a Sunday his is the one of two writers I always read first, the other being Janet Daley.” (source:UK-RW-POL:3)*

Example 7.5.26 *“I have always wondered how come he got the Chance to play for India, a very ordinary Street Level, even u-19 bowlers can Bowl at better pace than him.” (source:IND-RW-SPR)*

Example 7.5.27 *“I’ve always felt sorry for Laura Ingraham. Conservative pundits always live in fear that their ‘old ways’ are being threatened and live in fear of change.” (source:USA-LW-ENT)*

In the above examples from the three sub-corpora, the adverb ‘always’ was used to mean every time or all the time. In the first example from the UK sub-corpora, the commenter used ‘always’ to refer that every time he admired Christopher Booker

(a journalist) while reading the newspaper on Sunday. In general, the adverb ‘always’ falls into the category of adverb of frequency that indicates indefinite frequency. In this example, the adverb ‘always’ was qualifying the verb ‘admire’ and was used to refer to an unspecified amount of time. In the second example from the India sub-corpora, the commenter used the adverb ‘always’ to refer to Vinay Kumar (an Indian cricket player). Here, the commenter criticized the capability of Vinay Kumar as a bowler. In this example, the adverb ‘always’ qualifies the verb ‘wonder’ and it was used to mean every time or an unspecified amount of time. In the third comment from the USA sub-corpora, the commenter spoke about Laura Ingraham (an American conservative television host). The commenter mockingly criticized the conservatives here, and in this regard, he used the adverb ‘always’ to mean every time or all the time. Here, the adverb ‘always’ qualifies the verb ‘live’ and was used to refer to an undefined amount of time.

Adjectives: We noticed the common adjectives from the three sub-corpora were ‘clear’ and ‘vast’. The adjective ‘clear’ was used very less by the India commenters, while it was used almost more than 10 times in the UK and USA sub-corpora. The adjective ‘vast’ was used almost the same in India and the USA sub-corpora, while it was used by the UK commenters more frequently. We noticed that another adjective ‘substantial’ was not used for a single time in the UK sub-corpora; however, it was used by India and the USA commenters for 1.81 and 2.06 per 10,000 words, respectively. The other most frequently used adjectives are *thrilling*, *extraordinary*, and *evident*. We present here some examples of the adjective ‘clear’ from our corpus.

Example 7.5.28 “*Amuses me how remainers accuse the government of not having a plan for Brexit (which was true as they thought we would vote remain) and yet here they are blissfully avoiding what they would do if we again vote for a clear leave? It is exactly the same. And that is because they have no desire to respect democracy as to them it means ‘we tolerate your views and votes so long as you agree with us’. A tired, outdated record now.*” (*source:UK-LW-POL:2*)

Example 7.5.29 *“EVEN TODAY those arabs consider muslims of India-Pak-Bangla as their slaves.....Inspite of all this, converts STILL following their own tormentors is a clear case of Mental Slavery, which got ingrained due to prolonged inter-generational physical slavery.....While the compulsion to convert to save lives can be UNDERSTOOD, but foolishness to continue living as SYMBOL OF SLAVERY is outrightly wrong!” (source:IND-LW-POL)*

Example 7.5.30 *“When Colin Kaperich dressed for a media-attended practice by wearing socks visibly depicting police officers as pigs, he made his mission in the culture wars very clear. Michael Bennett was even less subtle when he recently sacked a white quarterback and immediately celebrated by giving the black power salute. The mainstream media almost uniformly gives these idiots condescending pats on the head-gold stars for being such good social justice warriors.” (source:USA-RW-POL:2)*

In the first example, the UK commenter used the adjective ‘clear’ to refer to something that has no doubt or something that is totally comprehensible. The commenter had doubts about the remainers’ (who did not support Brexit and wanted to remain in the EU) stance if the citizens of the UK get another chance to vote for Brexit. Here, the adjective ‘clear’ highlights the commenter’s confidence and refers to something that has no confusion. In the second example from the India sub-corpora, the commenter used the adjective ‘clear’ to refer to something that has no doubt. According to this commenter, Arab Muslims do not regard the converted Muslims (who belonged to some other religions earlier) from the Indian sub-continent (which includes India, Pakistan, and Bangladesh) and treat them as their slaves. However, as the Muslims of India, Pakistan, and Bangladesh still follow the Arabs (referred to as tormentors by the commenter), it was a transparent case of mental slavery, according to the commenter. Here, the adjective ‘clear’ was used to refer to the fact that it is understandable and transparent. By using ‘clear’ the commenter showed his confidence in the idea that he was talking about. In the third example from the USA sub-corpora, the commenter spoke about Colin Kaperich’s (an American football player and social activist) actions. Here

the adjective ‘clear’ was used by the commenter to show transparency to his statement. Also, it was used to show the emphasis and confidence of the commenter.

Phrases: Here, we would like to mention that only 4 booster-phrases were found in the whole corpus that we showed in table 7.4. The two common phrases from the three sub-corpora are ‘of course’ and ‘in fact’. The India commenters used ‘of course’ and ‘in fact’ less frequently than the UK and USA commenters. It was observed that two phrases, ‘for the most part’ and ‘in effect’ were used only by the USA commenters. Here, we show some examples of the phrase ‘of course’ from the UK, India, and the USA sub-corpora.

Example 7.5.31 *“Queered, how trendy, doesn’t matter that her music is garbage of course.”*
(*source:UK-LW-ENT:2*)

Example 7.5.32 *“Congrats Team India! & of course Saina & P V Sindhu-you make us all proud! 26 Golds for India, that is a fantastic achievement (11 more than ‘Kanada’, that’s almost 75% more-WOW !)...”* (*source:IND-LW-SPR*)

Example 7.5.33 *“I’m looking forward to male-born , high school seniors who suddenly discover that they’re women and demanding admission to all women’s colleges and bastions of feminism such as Smith, Barnard, and Mt.Holyoke. Of course, the-newly transgendered women don’t have to wear long hair, makeup, or dresses, which are social constructions anyway. Of course, the self-declared transgendered are under no legal obligation have to undergo sex change operations and hormone treatments. What helps save us is how the social revolutionaries always end up fighting among themselves.”*
(*source:USA-RW-SPR*)

In the examples mentioned above from the UK, India, and the USA sub-corpora, the commenter used the booster phrase ‘of course’ to mean something needless to say. In the first comment from the UK sub-corpora, the commenter criticized Lady Gaga (the famous American singer) and her music. It showed the commenter’s confidence

about something that was obvious to him, and he was trying to involve the readers in the conversation by showing his confidence. In the second comment from the India sub-corpora, the commenter used the phrase ‘of course’ to mean that it is needless to say the whole country is proud of the Indian players who won in the Commonwealth games, specifically P.V.Sindhu and Saina Nehwal (Indian Badminton players). In the third comment, the commenter talked about the self-declared transgenders who, despite having any social obligations to change their sex, want to have the facilities that are provided to women. In general, transgender refers to a person whose gender was different at the time of birth or during his/her upbringing and later transformed to some other gender. Here, the commenter criticized the trend where people unofficially declare themselves as transgenders and expects others to accept them as well. In this example, the phrase ‘of course’ meant something that is already known and, needless to say. It expresses the writer’s certainty about the particular propositions.

Comparison of the occurrences of *boosters*

From Figure 7.8, we extracted some notable characteristics of using *boosters* in the comments by the commenters of India, the UK, and the USA.

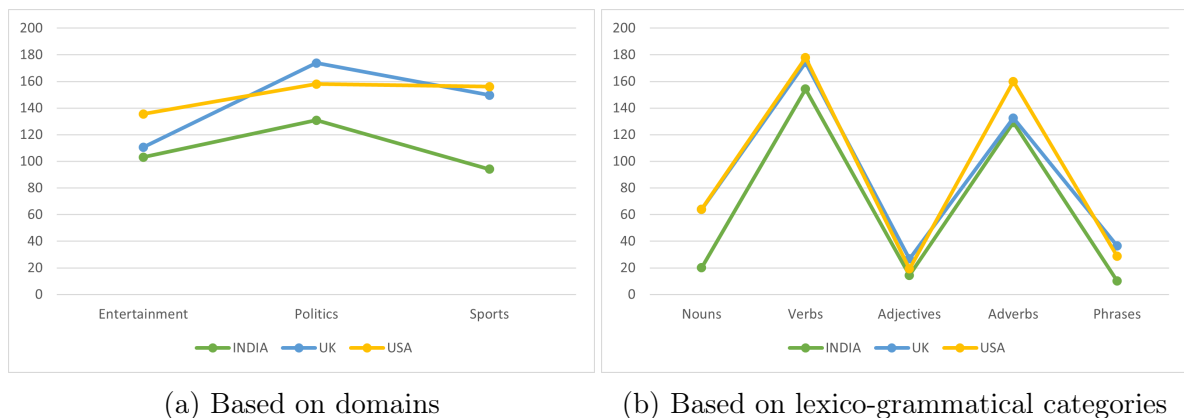


Figure 7.8: Comparisons of the occurrences (normalized) of *boosters*

- From Figure 7.8a, we can see that all the commenters, irrespective of countries, used *boosters* more frequently in the politics domain than the other two

domains, more specifically, the UK commenters employed them most compared to India and the USA commenters. The obtained frequency pattern with the three countries was almost identical. The commenters of all the countries, except India, used *boosters* less frequently in the entertainment domain. Notably, the USA commenters employed *boosters* almost the same in the politics and sports domain and a bit less in the entertainment domain which shows that the commenters of these domains were more confident about the propositions that they were claiming for.

- Analysing our corpus based on the lexico-grammatical categories of *boosters*, we conclude from the Figure 7.8b that booster-nouns, booster-adjectives, and booster-phrases were employed less frequently than booster-verbs and booster-adverbs by all the commenters, irrespective of three countries. Notably, the usage pattern of the lexico-grammatical categories of *boosters* was almost the same for the three nations. We noted differences in the usage (normalized) of lexico-grammatical items; more specifically, the USA commenters significantly used booster-adverbs than the India and UK commenters. Also, we would like to mention that booster-verbs were used almost the same by USA and UK commenters, whereas in the India sub-corpora, it was used a bit less which highlights the certainty and confidence of the UK and USA commenters.

7.5.3 Interactional: Attitude Markers

Another category of interactional metadiscourse markers is *attitude markers*. According to Hyland (2005b), *attitude markers* express the writer's affective and evaluative attitude towards propositions and convey their personal or professional feelings such as agreement, frustration, surprise, importance, etc. Writers use them to let their readers know about their feelings and opinions about a particular text. This category is used by the writers as a medium of communication with their readers where the writers convince the readers about their propositions. An investigation on digital news comments

written in Spanish was carried out by Moya Muñoz (2016) that reported significant occurrences of *attitude markers* in Spanish news comments. Analyzing our corpus, we confirmed that *attitude markers* played an indispensable role along with other markers in digital news comments written in English. In our study, *attitude markers* were used for the following purposes by the commenters: 1) to express the attitude and emotion of the commenters to particular propositions, 2) to express agreement or disagreement or judgment of the commenter about a proposition. Our results indicate that the commenters of India, the UK, and the USA used *attitude markers* most frequently in their comments. While comparing with previous studies, it was observed that similar to our study, in the studies of Hyland (2005a), Hyland (2005b), Mur-Dueñas (2011), Dueñas (2010), *attitude markers* were frequently used by the writers. On the contrary, in the studies of Cao and Hu (2014), and Yao (2022), *attitude markers* have the least frequency. In our study, a significant difference was also noticed in the frequency of using the lexico-grammatical categories of *attitude markers* among the three nations.

Table 7.5: Normalized frequencies of *attitude markers* found in India, the UK, and the USA

		Entertainment			Politics			Sports		
		LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
INDIA	Nouns	2.77	13.27	16.04	14.06	7.25	21.31	11.13	4.67	15.80
	Verbs	2.77	10.62	13.39	12.98	21.74	34.72	28.93	37.38	66.32
	Adjectives	130.27	148.62	278.89	73.54	126.83	200.37	171.38	271.03	442.41
	Adverbs	0.00	2.65	2.65	6.49	3.62	10.11	2.23	14.02	16.24
	Phrases	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total	135.81	175.16	310.97	107.06	159.45	266.51	213.67	327.10	540.77
UK	Nouns	6.67	0.00	6.67	8.16	14.97	23.13	7.97	7.96	15.93
	Verbs	6.67	15.40	22.06	15.16	16.38	31.53	21.92	2.65	24.57
	Adjectives	130.00	177.06	307.06	114.26	93.58	207.84	165.39	98.17	263.56
	Adverbs	3.33	0.00	3.33	9.33	3.74	13.07	2.99	0.00	2.99
	Phrases	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total	146.67	192.46	339.12	146.90	128.67	275.57	198.27	108.78	307.05
USA	Nouns	8.57	20.57	29.14	5.68	20.51	26.19	11.86	22.63	34.49
	Verbs	4.67	14.40	19.07	10.22	16.95	27.17	5.08	9.70	14.78
	Adjectives	101.25	78.16	179.41	111.32	129.78	241.10	111.81	110.99	222.80
	Adverbs	0.78	8.23	9.01	5.11	3.12	8.23	6.78	1.08	7.85
	Phrases	0.00	0.00	0.00	0.00	0.45	0.45	0.00	0.00	0.00
	Total	115.27	121.35	236.62	132.33	170.81	303.14	135.52	144.40	279.92

Table 7.5 summarize the occurrences (based on a normalized frequency of 10,000 words) of *attitude markers* in digital news comments by the commenters of the UK, India, and the USA). Here, we present and analyze the results based on domains as

well as political ideology. Irrespective of domains and political ideology, the following lexico-grammatical categories of *attitude markers* were frequently used by the commenters: nouns (*problem, wealth, limitation, value, absence, contribution, credibility, failure, importance, support, key, insight, shortcoming, complexity, hurdle, validity, well, consistency*), verbs (*contribute, deserve, ensure, extend, fail, ignore, lack, limitation, neglect, respond, support, overlook, expand*), adjectives (*only, better, new, major, good, hard, well, great, true, first, influential, important, core, poor, best, main, relevant, difficult, obvious, serious, interesting, worthwhile, acceptable, necessary, significant, central, dangerous, meaningful, easier, wise, useful, critical, satisfactory, worth, limited, complex, reasonable, valid, missing, short of, ripe, primary, tremendous, effective, surprising, notable, adequate, valuable, confident, promising, crucial, unfortunate, broad, narrow, inconsistent, comprehensive, consistent, unique, in-depth, essential, hopeful, valid, fundamental*), adverbs (*importantly, only, unfortunately, interestingly, positively, broadly*). Only one phrase (*go beyond*) was used by the USA commenters of the right-wing politics domain.

India: The quantitative analysis of the *attitude markers* used by the India commenters with respect to different domains shows that overall they used them more in the sports domain compared to the entertainment and politics domains. Among the lexico-grammatical categories, the usage of *attitude markers* by the India commenters of the sports domain shows that adjectives were mostly used, and there was not a single use of phrases by the India commenters. With respect to political ideology, the right-wing commenters of the sports domain used adjectives 1.58 times more in comparison to the left-wing. A significant difference was noticed in the use of adverbs where the commenters of the left-wing used 2.23 per 10,000 words and the right-wing commenters used 14.02 per 10,000 words which shows the unwillingness on the part of the commenter to show their attitudes towards their the other commenters. The India commenters of the entertainment domain used the lexico-grammatical categories of *attitude-markers* in the following order: adjectives, nouns, verbs, and adverbs. The

most frequently used category was adjectives (278.89 per 10,000 words), and phrases were not seen to be used for a single time, irrespective of political ideology. Also, we noticed that considering political ideology, the commenters of the right-wing sub-corpora used all the lexico-grammatical categories of *attitude markers* more than the left-wing. Table 7.5, shows that the commenters of the politics domain used the lexico-grammatical categories of *attitude markers* in the following order: adjectives, verbs, nouns, and adverbs. Taking into account the political ideology wise use, the right-wing commenters of the politics domain used adjectives and verbs more compared to the left-wing commenters. On the contrary, nouns and adverbs were used comparatively (almost 2 times) more by the left-wing commenters compared to the right-wing ones which suggests that the left-wing commenters were more eager to show their attitudes towards the other commenters.

The UK: The quantitative analysis of the UK comments shows that the commenters of the entertainment domain used *attitude markers* most frequently compared to the sports and politics domains. The UK commenters of entertainment, sports, and politics domains used the lexico-grammatical categories of *attitude markers* in the following order: adjectives, verbs, nouns, and adverbs. Phrases were not used for a single time by the commenters of the UK, irrespective of domains and political ideology. In the entertainment domain, a difference was noticed in the use of verbs and adjectives where the right-wing commenters used the verbs and adjectives comparatively more than the left-wing. Also, while the left-wing commenters used nouns and adverbs for 6.67 and 3.33 per 10,000 words, respectively, not a single use of nouns and adverbs was seen by the right-wing commenters of the entertainment domain. As a possible reason for this could be the right-wing commenters of the UK were not so eager to show their affective values to the other commenters. In the sports domain, taking into consideration political ideology, it was noticed that the commenters of the left-wing used all the lexico-grammatical categories of *attitude markers* more than the right-wing. A remarkable difference was noticed in the use of verbs where the right-wing commenters used

verbs for 2.65 per 10,000 words, and the left-wing commenters used it more frequently (21.92 per 10,000 words), which is almost 10 times more. However, the use of nouns was almost the same in the left-wing and right-wing sub-corpora. Also, the left-wing commenters used adverbs for 2.99 per 10,000 words, whereas not a single use of adverbs was noticed among the right-wing commenters. In the politics domain, the left-wing commenters used adjectives and adverbs more in comparison to the right-wing while the right-wing commenters used nouns and verbs more than the left-wing (See Table 7.5). In the case of adverbs, it was observed that the left-wing commenters used them three times more than the right-wing ones. The use of verbs was almost the same for left-wing and right-wing, while in the use of adjectives, the left-wing commenters were seen to use comparatively more than the right-wing.

The USA: The quantitative analysis of the USA comments shows that the commenters of the politics domain used the lexico-grammatical categories more, followed by the sports and entertainment domains. The USA commenters of the politics domain used the lexico-grammatical categories of *attitude markers* in the following order: adjectives, verbs, nouns, and adverbs. In relation to political biasness, it was noticed that all the lexico-grammatical categories, including phrases, were used by the right-wing commenters. In contrast, the left-wing commenters used all the markers except phrases. Irrespective of all the domains and political ideology, phrases were seen to be used only by the right-wing commenters of the politics domain (0.45 per 10,000 words). The right-wing commenters used nouns, verbs, and adjectives more; on the other hand, adverbs were seen to be used a bit more by the left-wing commenters. In the sports domain, the right-wing commenters used nouns and verbs more, while adjectives and adverbs were used more by the left-wing. Although adjectives were almost similar in left-wing and right-wing, noticeable differences were observed in the use of adverbs, verbs, and nouns. While the right-wing commenters used adverbs less frequently, the left-wing used them almost 5 times more which depicts the inclination of the left-wing commenters in showing their attitudes and judgements to the readers

and other commenters. On the contrary, the right-wing commenters used nouns and verbs more frequently than the left-wing. The right-wing commenters of the entertainment domain used nouns, verbs, and adverbs more than the left-wing; however, the left-wing commenters used adjectives more than the right-wing. The right-wing commenters used adverbs significantly more (8.23 per 10,000 words) than the left-wing (0.78 per 10,000 words).

Occurrences in Lexico-Grammatical categories

Table 7.6 depicts the five most frequent (normalized frequency per 10,000 words) occurrences of the lexico-grammatical categories of *attitude markers* (verbs, nouns, adjectives, and adverbs) by the commenters of India, the UK, and the USA.

Table 7.6: Top-5 (normalized frequency per 10,000 words) words of lexico-grammatical categories of *attitude markers*

	INDIA		UK		USA	
Verb	support	57.80	support	26.83	support	14.30
	fail	21.70	fail	12.92	fail	12.93
	deserve	18.10	deserve	11.10	ignore	12.44
	lack	5.75	ignore	8.27	lack	9.05
	ignore	3.74	lack	7.02	extend	6.42
Noun	problem	21.26	problem	14.23	problem	47.43
	value	10.87	value	8.81	support	14.14
	support	7.96	failure	5.33	failure	7.16
	failure	4.82	key	5.11	value	6.17
	importance	3.85	wealth	3.77	wealth	3.62
Adverb	only	16.18	only	14.86	unfortunately	11.56
	unfortunately	9.94	unfortunately	2.40	only	6.27
	importantly	2.89	importantly	2.13	importantly	3.75
					interestingly	2.06
					positively	1.01
Adjectives	great	177.18	good	109.00	good	107.67
	only	145.14	only	102.18	only	90.15
	good	132.90	great	73.42	well	52.58
	well	95.21	well	70.17	better	51.29
	best	73.08	best	52.90	great	47.94
Phrases					go beyond	0.45

Verbs: We got 13 attitudinal verbs used in the UK, India, and the USA sub-corpora. From Table 7.6, we can see that among the verbs, ‘support’ and ‘fail’ were most frequently used by the commenters of all three nations; however, these were most frequently used by the India commenters (57.80 and 21.70 respectively per 10,000

words). Among the other two nations (the UK and the USA), the UK commenters used the verbs ‘support’ and ‘fail’ more than the USA commenters. Our study coincides with the study of Dueñas (2010), where the mostly used attitudinal verb was ‘support’. Also, another two verbs (‘lack’ and ‘ignore’) were found to be used by all the commenters, particularly, the USA commenters used them a bit more than the UK and India commenters. Another verb ‘deserve’ was found to be frequent in the India and UK sub-corpora, whereas in the USA sub-corpora, the verb ‘extend’ was found to be more frequent. Here we present some examples of the verb ‘support’ from the UK, India, and the USA sub-corpora.

Example 7.5.34 *“I love the passion of the article and it warms my heart that Lady GaGa was a core support to Brian and many others. As a straight man and a music fan I never much cared for her music or image, give or take a couple of really decent pop songs, but that’s really not the point.” (source:UK-LW-ENT-2)*

In this example from the UK sub-corpora, the commenter used the attitudinal verb ‘support’ to mean to help or assist something or someone. The commenter described that he was overwhelmed to see that Lady Gaga (a famous American singer and songwriter) came in support of Brian (an American bandleader) and many others while no other celebrities came in support of them. Here, the ‘support’ is used by the commenter to mean Lady Gaga’s help or assistance towards her fellow music artists. It was used in a positive sense.

Example 7.5.35 *“just see nonsense. wake up countrymen. Trust the best PM since Independence and one the great Statesman to decide the best for Mother India. Please trust me Modi is gift to India. Leader’s like him are born only once in a blue moon. Support him when is as life for moksa. he is avatar.” (source:IND-RW-POL)*

In the India sub-corpora, the attitudinal verb ‘support’ was used by the commenter while pleading to the fellow citizens of India to have their faith in the present Indian

Prime Minister, Shri Narendra Modi. Here, ‘support’ was used to mean ‘in favor of’, and it was used in a positive sense. By using this verb, the commenter was showing his complete faith in the efficiency of the present Prime Minister of India. The commenter here was trying to convince the readers by showing his attitude and also including them in the conversation.

Example 7.5.36 *“I don’t think its cruel or ineffective and I stand by in support of POTUS Trump. We aren’t talking about local drug dealers! We are talking about the international drug trade. In some cases these drug cartels are so powerful that they are given safe haven because they control the nation they reside and they are powerful enough to other countries with drugs. Yes the British flooded China with Opium as did the French in China and Indochina . Vast swaths of population addicted to drugs. Yes there were popular uprisings at the time to stop the addiction and exploitation of their citizenry but their uprisings failed. Today we are faced with the same question? Whose civil liberties are we as a nation going to protect?” (source:USA-RW-POL-1)*

In the USA sub-corpora, the commenter used the attitudinal verb ‘support’ to mean ‘in favor of something’. Here, the commenter was in favor of Donald Trump (the former President of the USA) and the President’s decision against the drug dealers. Later, he described various incidents caused due to drugs and their adverse effect on mankind. The commenter used the verb ‘support’ in a positive sense. Like the previous examples, the commenter’s attitude was evident while using the verb ‘support’. This also highlights the commenter’s endeavour to include the readers as well as the other commenters in the conversation.

Nouns: We found that from our corpus, some of the nouns were used notably more, and some were used significantly less. If we compare our study with Carrió-Pastor (2019a) we can see that in our study, attitudinal nouns are the third most frequently used attitude marker category for UK and India, whereas attitudinal nouns are the second most frequent category in Carrió-Pastor (2019a). Regarding the nouns found in our corpus, it was evident from Table 7.6 that although the frequency of using the

three nouns ‘problem’, ‘failure’, and ‘value’ differs, they are used by the commenters of all three nations. The noun ‘problem’ was used most frequently by the UK, India, and USA commenters. While in the India and UK sub-corpora, ‘value’ was counted among the first two mostly used nouns (10.87 and 8.81 per 10,000 words), in the USA sub-corpora, it was used in fourth place (6.17 per 10,000 words). Here we show some examples of the attitudinal noun ‘value’ from the UK, India, and the USA sub-corpora, which were used by the commenters in three different senses.

Example 7.5.37 *“Head in sand, delusional nonsense. Have you seen the value of the pound and the performance of our economy you muppet-project fear is project reality, staring you in the face, but still you can’t see it.” (source:UK-RW-POL-3)*

Example 7.5.38 *“Very less comments cause its not BL00DY cricket.. shame on people who don’t value achievements of these players.. Congratulations to all and wish you all the best..” (source:IND-RW-SPR)*

Example 7.5.39 *“My mother always said “Jimmy Carter was too honest to be President”. I have always agreed. He is a man that has lived by his own moral code with unapproachable value and integrity.” (source:USA-LW-POL)*

In the first example from the UK sub-corpora, the commenter used the attitudinal noun ‘value’ to mean worth or merit. Here, the commenter expressed his views regarding the effects of Brexit on the economic situation of the UK. He criticized the present situation that affected the currency and the economy of the UK. According to the study of Dueñas (2010), the noun ‘value’ falls under the category of *attitude markers expressing assessment*. Following that in this example, the commenter was using the noun ‘value’ for assessing the worth of the currency of the UK. By using this noun, he was including the readers and other commenters in the conversation. In the second example from the India sub-corpora, the commenter used the attitudinal noun ‘value’ to acknowledge something. He used the noun while criticizing the other commenters

who do not acknowledge the achievements of the other players other than the cricketers. Here it was used to mean to regard something. In the third example from the USA sub-corpora, the commenter used the attitudinal noun ‘value’ to refer to ethical values and principles. Here, the commenter talked about former American President Jimmy Carter, who, according to the commenter was an honest man with ethical values and principles and never compromised with his principles. In this comment, ‘value’ was used in a positive sense which means merit. So, in all the above examples, the commenters used the attitudinal noun ‘value’ to show their attitude about something and to include the readers in the conversation.

Adverbs: Adverbs was the fourth most frequent category used by the commenters of the UK, India, and USA sub-corpora. Our study inlines with the study of Dueñas (2010) where adverb comes in the fourth place. There were three adverbs that were used by the commenters of all three nations: ‘only’, ‘unfortunately’, and ‘importantly’, though the frequency of using them was not the same. Another two adverbs, ‘interestingly’ and ‘positively’ were used less frequently only in the USA sub-corpora. Here, we illustrate some examples of the adverb ‘only’ which were extracted from our corpus.

Example 7.5.40 *“Almost half of us in the UK agree that a new nationwide referendum should be held, this time on the terms for leaving the EU (survey reported January 2018). Far fewer, only a third of those questioned, are against the idea.” (source:UK-LW-POL:2)*

Example 7.5.41 *“Only 10-15 of our people would always prefer Congress Govt should come to power so that they can carry out tax evasion/corrupt activities. Further Corrupt Sarkari Babus always prefer Congress Party to rule states and Centre. only sufferer are Aam Aadmi/Middle/Lower/Middle and working class..” (source:IND-LW-POL)*

Example 7.5.42 *“This is always true in any business. The clients who give you the most business get the best deals. There are probably only a few items in the average American household that haven’t been made in China.” (source:USA-LW-POL)*

In the first example UK sub-corpora, the adverb ‘only’ was used by the commenter to mean something low in number or amount. Here, the commenter expressed his opinion regarding the need for a second referendum on Brexit. The adverb ‘only’ was used to refer to a small number of people who opposed the idea of the second referendum. It was used in a positive sense by the commenter. By using ‘only,’ he expressed his judgement and showed his attitude to the other commenters. In the second example from the India sub-corpora, the commenter used the adverb ‘only’ to indicate something of a low number or amount. The commenter used ‘only’ to mean that a small number of people out of the whole population of India would prefer Congress Government so that they could continue corruption in the country. He criticized the government employees who are supporters of the Congress Party and enjoys various corrupt activities, and according to the commenter, the only sufferer is the middle-class people. By using ‘only’, the commenter showed his attitude towards the proposition. Like the examples mentioned above, in the third example from the USA sub-corpora also, the commenter used the attitudinal adverb ‘only’ to indicate something low in number. He stressed the fact that it is very likely that the clients who are giving more business to a company will have the benefits and the best deals as well. It was used by the commenter in a positive sense. Also, it showed the attitude of the commenter.

Adjectives: Among the adjectives, the commenters of the UK, India, and the USA used ‘great’, ‘only’, ‘good’, and ‘well’ with different frequencies. These adjectives were more frequently used by the India commenters, followed by the UK and the USA commenters. Apart from these adjectives, ‘best’ (in the India and UK sub-corpora) and ‘better’ (in the USA sub-corpora) were used frequently. The less frequent adjectives used in the UK and India sub-corpora are *unique, in-depth, essential, hopeful, valid, and fundamental* which were not used by the UK and India commenters for a single time. However, in the USA sub-corpora, these adjectives were among the least frequent ones. Some examples of the adjective ‘good’ from the UK, India, and the USA sub-corpora are given below:

Example 7.5.43 *“Well, this may be more heartfelt than that utterly appalling ‘Best Female Rapper’ feature but I’m afraid it doesn’t stand up to much scrutiny. GaGa may be good at publicising herself, but she doesn’t really have much to offer musically that hasn’t been done before. Kylie, Cher, Madonna and Diana Ross all championed LGBT issues in a less ‘LOOK AT ME !!!’ way and there have successor then to Sir Noel Coward.” (source: LW-UK-ENT:2)*

In this example from the UK sub-corpora, the commenter spoke about Lady Gaga (a famous American singer and songwriter) where he used the evaluative adjective ‘good’. Although he was not so fond of Lady Gaga, he believed that Gaga had a superior quality of publishing herself. According to the commenter, there were various previous artists who were talented enough, even better than Gaga, but unlike Gaga, they were not so eager to publish themselves. Here, ‘good’ is used in a general sense to mean something superior but not the best. Here, the implication of the evaluative adjective ‘good’ is in a positive manner. Following the categorization of *attitude markers* by Dueñas (2010), the evaluative adjective ‘good’ was used in this example as an attitude marker expressing assessment.

Example 7.5.44 *“Hockey India needs to look from inside out - - -not inside out - - -The forward line in talented Akashdeep Singh, Mandeep Singh, Gurjant Singh, SV Sunil and Dilpreet Singh is quite a potent entity. Add Ramandeep - - who in spite of his odd and unathletic posture - - is still good as roll in substitute. The defense line has the very talented Rupinderpal Singh and Harmanpreet Singh.” (source:IND-LW-SPR-1)*

In the India sub-corpora, here the commenter talked about the Indian Hockey player Ramandeep where he mentioned the evaluative adjective ‘good’. Here, it was used to mean a high or superior quality. According to him, Ramandeep, despite lacking supreme athletic qualities like the other talented players, should be placed as a substitute player in the forward line with the other talented hockey players. The commenter

used the adjective in a positive way, and like the previous example, it falls under the category of attitude marker expressing assessment (Dueñas (2010)).

Example 7.5.45 *“Believe it or not, not everyone is like you, some people really care about the society that they live in and common decency is part of their make-up. I guess people like you and Ingraham exist too, but thank goodness there are more good people than bad.” (source:USA-LW-ENT-1)*

In the USA sub-corpora, the commenter used the evaluative adjective ‘good’ to refer to a superior quality that people must own as we live in a civilized society. Here ‘good’ was used by the commenter to criticize the other commenter, who was insensitive as a human being and did not possess any righteous quality. Although it was used to criticize, it was used in a positive way. By using this adjective, the commenter expressed his attitude and judgement.

Phrases: Phrases were the least frequent category of *attitude markers* expressing attitudinal values occurred in our corpus. Only in the USA sub-corpora, the attitudinal phrase ‘go beyond’ was used. The UK and India commenters never used it once in their comments. Here, we mention one example of the phrase ‘go beyond’ from the USA sub-corpora.

Example 7.5.46 *“I could burn my draft card in 1969, but I never could bring myself to desecrate the flag. One action was against an out-of-control government, engaged in an unjustified war. The flag, the national anthem, go beyond the actions of our government; they are about who we are as a people, where we’ve been, and where we will go. They are symbolic of the hopes and dreams we all share as Americans. We should protest more, especially our unending wars.” (source:USA-RW-POL:1)*

In the USA sub-corpora, the commenter used the attitudinal phrase ‘go beyond’ to mean pass ahead of. Here, the commenter described the importance of the flag and the national anthem of the USA, which surpasses the actions of the government

and represents the voice of the common people. The commenter used the phrase ‘go beyond’ to criticize the deeds of the Government and to show the significance of the national flag and anthem. by using the phrase, the commenter here expressed his attitude towards the other commenters.

Comparison of the occurrences of *attitude markers*

Figure 7.9a depicts the interesting features of using *attitude markers* by the UK, India, and the USA commenters.

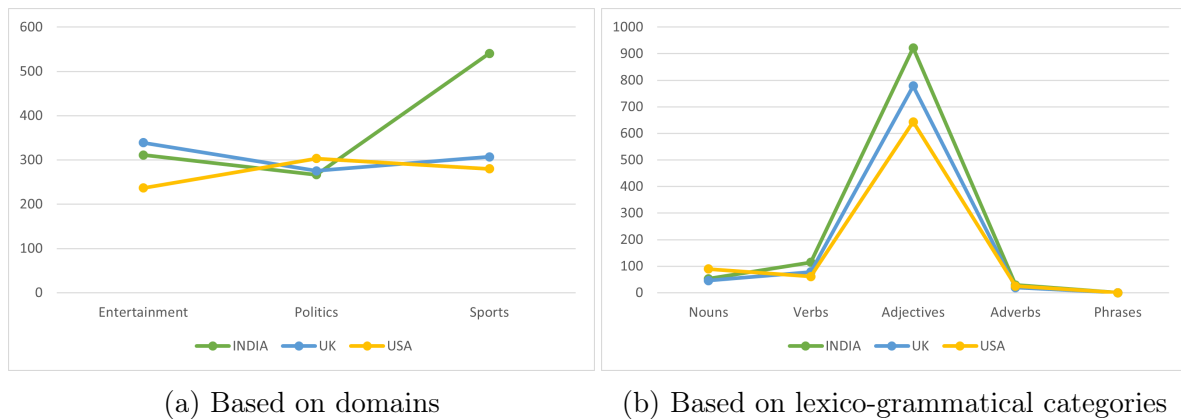


Figure 7.9: Comparisons of the occurrences (normalized) of *attitude markers*

- Figure 7.9a portrays that the India and UK commenters used *attitude markers* almost the same in the entertainment and politics domains, whereas the India commenters used them mostly in the sports domain. Contrarily, the USA commenters of the politics domain used them more in comparison to the entertainment and sports domains. The possible explanation for the differences in the frequency could be that the India commenters of the sports domain argue more with the other commenters, and the USA commenters of the politics domain contradict more than the other two nations. Also, they display their attitudes openly without restricting them to their readers and other commenters.
- Regarding the lexico-grammatical categories of *attitude-markers*, we can see that the pattern of using them was the same in all three nations. The India com-

menters used adjectives most, following the UK and the USA commenters. The least frequently used *attitude marker* was phrases, as they were not used for a single time in the India and the UK sub-corpora, whereas they were used slightly in the USA sub-corpora.

- Figure 7.9a portrays that the India and the UK commenters used *attitude markers* almost the same in the entertainment and politics domains, whereas the India commenters were seen to use them mostly in the sports domain. Contrarily, the USA commenters of the politics domain used *attitude markers* more in comparison to the entertainment and sports domains. The possible explanation for the differences in the frequency could be that the India commenters of the sports domain and the USA commenters of the politics domain were more eager to express their thoughts and judgements in front of the other commenters.

7.5.4 Interactional: Engagement markers

Whereas the above-mentioned metadiscourse markers express the author's stance about a proposition, the *engagement marker* is one of the interactional markers used by the writers to engage the readers as participants within the text. The writers use explicit or implicit ways to involve the readers as discourse participants (Hyland (2005b)). They use them to establish a relationship with the readers. Previous studies, such as, Hyland (2005a), Hyland (2005b), Mameghani and Ebrahimi (2017), have investigated the use of *engagement markers*. In our study, *engagement markers* were used in the following ways: 1) to express the commenter's point of view to the readers (in our case, to other commenters or readers) about a particular proposition where the writer includes the readers as participants in the conversation. 2) to express the politeness of the commenter towards another commenter or the reader. 3) to build a relationship with the readers by introducing their voices. We found that in our study, the commenters used *engagement markers* adequately while commenting on digital news, which contrasts with the study of Mur-Dueñas (2011), where the writers used fewer *engagement mark-*

ers compared to the other interactional markers. The lexico-grammatical categories of *engagement markers* are personal references, questions, imperatives, and directives that a writer uses for reader participation. Table 7.7 outlines the normalized frequencies of *engagement markers* found in the India, UK, and USA sub-corpora. This section describes the in-depth analysis of the *engagement markers* based on domains and political ideology. We have not found the lexico-grammatical category named ‘questions’ in our corpus. The following lexico-grammatical categories of *engagement markers* were

Table 7.7: Normalized frequencies of *engagement markers* found in India, the UK, and the USA

		Entertainment			Politics			Sports		
		LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
INDIA	Personal references	138.58	71.66	210.24	48.66	94.22	142.88	60.09	51.40	111.50
	Questions	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Imperatives	11.09	0.00	11.09	6.49	3.62	10.11	4.45	4.67	9.12
	Directives	5.54	29.19	34.74	18.38	21.74	40.13	15.58	18.69	34.27
	Total	155.21	100.85	256.06	73.54	119.59	193.12	80.12	74.77	154.89
UK	Personal references	83.33	103.93	187.26	176.05	142.24	318.29	126.53	82.25	208.78
	Questions	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Imperatives	3.33	0.00	3.33	7.00	1.40	8.40	5.98	7.96	13.94
	Directives	13.33	23.09	36.43	26.82	19.18	46.00	14.94	15.92	30.86
	Total	100.00	127.02	227.02	209.86	162.82	372.69	147.45	106.13	253.58
USA	Personal references	40.50	30.85	71.35	28.97	33.89	62.86	27.10	32.33	59.43
	Questions	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Imperatives	0.00	2.06	2.06	2.27	7.58	9.85	5.08	3.23	8.31
	Directives	19.47	10.28	29.76	9.66	7.58	17.24	22.02	10.78	32.80
	Total	59.97	43.19	103.17	40.89	49.06	89.95	54.21	46.34	100.55

used by the commenters of India, the UK, and the USA irrespective of domains and political ideology: personal reference (one, me, you, us, our, we), imperatives (consider, see, notice, want, feel, suppose, note that) and directives (need to, should, must).

India: If we focus on table 7.7, we can see that the India commenters of the entertainment domain have overall used *engagement markers* more frequently, followed by the politics and sports domains. The mostly used lexico-grammatical category, irrespective of domains and political ideology, was the personal reference. Concerning political ideology, in the entertainment domain, the left-wing commenters used overall *engagement markers* more than the right-wing. The left-wing commenters used personal references more frequently than the right-wing, more than 2 times. In the use of imperatives, not a single imperative was used by the right-wing commenters of the

entertainment domain, whereas the left-wing used it for 11.09 per 10,000 words which points out the fact that the left-wing commenters were more eager to include the readers as participants in the conversation. On the contrary, the right-wing commenters used directives more (almost 6 times more) compared to the left-wing. In the politics domain, the overall use of *engagement markers* was seen more in the right-wing sub-corpora. Personal references and directives were used more by the right-wing, whereas imperatives were used more by the left-wing commenters (almost 2 times more). In the sports domain, personal references were used more by left-wing, and directives were used comparatively more by the right-wing. A possible reason could be that the left-wing commenters of the sports domain were more eager to bring the readers or other commenters explicitly into the conversation. and the right-wing commenters were more determined to instruct the other commenters and readers about certain propositions. The use of imperatives was almost the same in both the left-wing and right-wing sub-corpora which indicates the tendency of the commenters of both the left-wing and right-wing to consider and regard the readers as participants in the conversation.

The UK: Regarding the domains, the UK commenters used the lexico-grammatical categories of *engagement markers* overall more in the politics domain, followed by the sports and entertainment domains. In the politics domain, personal references and directives were used more frequently by left-wing commenters compared to right-wing ones. Also, imperatives were used almost 6 times more in the left-wing, whereas they were used very low in the right-wing sub-corpora. The same reason, as in the case of India, could be thought that the left-wing commenters of the politics domain were inclined more to include the readers as participants in the conversation. In the sports domain, personal references were used much more in the left-wing sub-corpora than in the right-wing. On the other hand, imperatives and directives were used comparatively more by the right-wing commenters than the left-wing; however, the difference was not so noteworthy. In the entertainment domain, the most frequently used category was personal reference which was used mostly by right-wing commenters. The next

most frequently used category was directives which were also used more by right-wing commenters. However, there was not a single use of imperatives by the right-wing commenters, whereas it was used on a lower side by the left-wing commenters (3.33 per 10,000 words) which suggests the disinterestedness of both the right-wing and left-wing commenters of the entertainment domain in including the readers in the conversation.

The USA: Regarding domains, the commenters of the entertainment domain overall used the lexico-grammatical categories of *engagement markers* more, followed by the sports and politics domain. In the entertainment domain, left-wing commenters used personal references and directives more than right-wing ones. However, the left-wing commenters did not use imperatives a single time, whereas it was used a bit in the right-wing sub-corpora (2.06 per 10,000 words). The most frequently used category in the sports domain, irrespective of political ideology, was personal references, and they were used more by the right-wing commenters compared to the left-wing. On the contrary, imperatives and directives were used more frequently in the left-wing sub-corpora than in the right-wing. In the politics domain, the most frequently used category was personal references, which the right-wing used comparatively more. The left-wing commenters used imperatives comparatively less than the right-wing (almost 3 times more). Directives were used a bit more by the left-wing than the right-wing which suggests the instructive tendency of the left-wing commenters who were more eager to instruct their readers or the other commenters about performing an action.

Occurrences in Lexico-Grammatical categories

Unlike the previous metadiscourse categories, in the case of *engagement markers*, as the number of occurrences was on the lower side in our corpus, we decided to put all of them in the table 7.8. Here, we show the results from India, the UK, and the USA.

Personal references: Personal references refer to the personal pronouns and the reader pronouns that the writers use to include the readers in the texts by sharing their viewpoints. The commenters of the three nations used six personal references ('you',

Table 7.8: Top-5 (normalized frequency per 10,000 words) words of lexico-grammatical categories of *engagement markers*

	INDIA		UK		USA	
Personal ref.	you	133.99	you	205.01	me	67.55
	us	116.16	we	175.71	one	52.03
	one	76.64	me	122.26	us	49.55
	we	52.49	one	94.71	you	19.23
	me	37.97	us	64.15	our	1.46
	our	36.21	our	52.49	we	0.45
Imperatives	see	16.21	consider	13.72	consider	14.96
	consider	10.14	see	10.95	note that	2.23
	notice	1.81	notice	1.00	suppose	1.69
	want	1.08			notice	0.89
	feel	1.08			see	0.45
Directives	must	59.45	need to	54.20	need to	54.08
	need to	44.68	must	52.29	must	25.71
	should	5.00	should	6.81		
Questions						

‘us’, ‘one’, ‘we’, ‘me’, and ‘our’); however, there was a difference in the frequency of using them. The personal pronoun ‘you’, and the indefinite pronoun ‘one’ were used comparatively more in the UK sub-corpora followed by the India and USA sub-corpora, whereas the personal pronoun ‘us’ was used more by the India commenters followed by the UK and USA commenters. From table 7.8, some notable differences were noticed in the use of ‘you’. The USA commenters used ‘you’ very less (19.23 per 10,000 words), whereas, the India and UK commenters used it significantly more (133.99 and 205.01 per 10,000 words, respectively). As a potential reason, it could be stated that the UK and India commenters were more inclined to acknowledge the reader’s presence in the conversation by using the personal pronoun ‘you’. Here, we discuss some examples of the personal pronoun ‘you’ from our corpus.

Example 7.5.47 *“I remember dancing along to her songs as a teenager with such euphoria. 10 years later, I’m in such a shit place but playing those same songs still brings a little bit of light to the crappy days and that’s all you really need from music innit? Cheers Gaga!” (source:UK-LW-ENT:2)*

In this comment from the UK sub-corpora, the commenter described his fondness for Lady Gaga’s songs during his teenage, and even after so many years, he loves to

listen to her songs. By using the personal pronoun ‘you’, the commenter includes the readers or the other commenters in the conversation along with him. Here, the commenter convinced the other commenters about his feelings regarding Lady Gaga’s songs and built a kind of rapport with them by including them as well.

Example 7.5.48 *“When money & success hits you high suddenly, the downfall will also be the same. His stupid tweet to PM, that very day his downfall was certain. not because politics or Modi, he has shown his arrogance to the highest office of the nation. No seasoned actors have dared to the nation. No seasoned actors have dared to write such tweet.” (source:IND-RW-ENT)*

In this example from the India sub-corpora, the commenter talked about the famous Indian Comedian Kapil Sharma, who, in an intoxicated state, showed his arrogance in a tweet to the present Prime Minister of India. Here, the commenter criticized the reckless behaviour of the comedian. By using the personal pronoun ‘you’, the commenter includes the readers and creates oneness with them. It includes his fellow commenter with whom he was having the conversation, the readers or other commenters who are not included in the conversation but can read the comments, and also the commenter could refer to the comedian Kapil Sharma himself.

Example 7.5.49 *“I’ll tell you something else they remain silent on, the black children in urban areas. If I were a black person in lets just say, Chicago, I would want to know why I wasn’t important to “celebrities”. Oh yeah, they live in rahm’s gun-controlled utopia.” (source:USA-RW-ENT)*

In this comment from the USA sub-corpora, the commenter questioned the importance of black children to celebrities. He actually criticized the attitude of celebrities towards black children. The personal pronoun ‘you’ was used by the commenter to include the readers and other commenters in the conversation. It highlighted the inclusive tendency of the commenter where he acknowledged the presence of other commenters.

Imperatives: Imperatives refer to the actions that the writers use to direct their readers about a particular proposition, also accepting their points of view. Among all the imperatives found in the corpus, ‘see’, ‘consider’, and ‘notice’ were the most common ones that occurred in all the three sub-corpora; however, there was a difference in the frequency of use. The other imperatives used by the India and USA sub-corpora were *want*, *feel*, *note that* and *suppose*, while these imperatives were not used for a single time in the UK sub-corpora which indicates UK commenters’ tendency to restrict themselves in the use of different imperatives as they do not want to raise a commanding tone to their readers or other commenters. In the use of the imperative ‘notice’, we found that it was used significantly less in all the three sub-corpora. However, ‘consider’ was used in a considerable way in all three sub-corpora, and ‘see’ was used remarkably less in the USA sub-corpora (0.45 per 10,000 words) compared to the India and UK sub-corpora. Here, we show some examples of ‘consider’ from our corpus.

Example 7.5.50 *“The difference with the N word, which is specific , is that queer and gay are appropriations from the common lexicon that mean very different things to the current meaning that can be applied to them. Would you consider these words to be off limits to the current population yet are freely spoken, like the N word, within a community in which they have their new meaning.” (source:LW-UK-ENT:2)*

In this comment from the UK sub-corpora, the commenter discussed the difference in using the N-word (which means nigger and refers to black people) with the terms gay and queer. According to him, the N-word and the terms gay and queer are used differently, where the former is used as an ethnic slur to a particular community of people, and the latter ones are used in a broader sense with no ethnic boundaries. Here, the commenter used ‘consider’ to mean thinking about something. By using ‘consider’, he includes the other commenters and the readers in the conversation.

Example 7.5.51 *“Now is the time when govt should consider sports as industry & invest at least Rs.10 thousand cr. every year on developing infrastructure. Catch them*

young should be the slogan which every state should follow. We should be among 10 top sporting nations in the world.” (source:IND-RW-SPR)

In this comment from the India sub-corpora, the commenter described the importance of investing in sports and Indian players. He advised the Indian government to think of sports as an industry. Here ‘consider’ refers to thinking about something which was used by the commenter to include the readers and other commenters in the conversation. Also, it was used to introduce the voice of the readers or other commenters.

Example 7.5.52 *“Laura Ingraham took it just a little too far. Her hatred of everything she considers not conservative enough got the best of her. David Hogg and his fellow classmates are the most brave thing I have seen in a long time. They are saying they have had enough.” (source:USA-LW-ENT)*

The USA commenter in this example narrated that Laura Ingraham (an American conservative television host) was rightly confronted by David Hogg (an American gun control activist famous for several high-profile protests, marches, and boycotts) for her highly conservative thoughts. Here, ‘consider’ was used by the commenter to refer to something that Laura Ingraham thinks about. However, it also represents the commenter’s (who is commenting) as well as the readers’ (in our case, the other commenters’) voice.

Directives: Directives refer to the ways in which the writer determines the actions of the readers by instructing them. We found three common directives, among which the obligation modals, such as ‘need to’ and ‘must’ were used in India, the UK, and the USA sub-corpora. Though with a variation in the usage, these directives were used significantly by the commenters of three nations. Another obligation modal, ‘should’ was used significantly less by the India and UK commenters, however, it was not used for a single time by the USA commenters.

Example 7.5.53 *“I am not sure if these people will EVER stop bleating, we need to get on with making it work. I am thoroughly ashamed of half of this country and their ‘can’t do’ attitude. We CAN make Brexit a success and we WILL do. Get behind YOUR country and stop this EU cult worship.” (source:UK-RW-POL:3)*

In this comment from the UK sub-corpora, the commenter criticized the immature behaviour of the people of the UK who think that they cannot do anything without the EU. He encouraged the people of the UK to make Brexit a considerable success. The obligation modal ‘need to’ refers to something that should be done or followed. Here, by using ‘need to’, the commenter engaged the readers and commenters in the discourse. Also, he used the inclusive pronoun ‘we’ before the obligation modal ‘need to’, which suggests the collective endeavour that, according to the commenter, the people of the UK could do and make Brexit successful.

Example 7.5.54 *“Every one needs to understand the health condition of Kapil. In that condition anyone would use foul language. kapil has abused himself as well. No one in the right frame of mind be self destructive. Please please give him space. It is heart breaking to see such talent getting wasted.” (source:IND-RW-ENT)*

The commenter of the Indian sub-corpora used ‘needs to’ to get the reader and other commenters to consider something. Here, the commenter talked about the famous Indian comedian Kapil Sharma. He showed concern about the comedian, who was not well mentally and physically. By using the obligation modal ‘needs to’, the commenter included the other commenters in the discourse. He expressed his concern in the form of a message to the readers and other commenters and convinced them to do what was necessary.

Example 7.5.55 *“China has out smarted America time and again. They took advantage of short sighted businesses. The race to the bottom that has left America weak and economically divided. While China has plotted and planned their economy.*

Americans need to stop name calling each other and pull together before it is too late.”
(source:USA-LW-POL)

In this example from the USA sub-corpora, the commenter cautioned about how China outsmarted the Americans and made them economically weak and divided. Here the obligation modal ‘need to’ was used by the commenter to advise his fellow countrymen to join hands together and work to strengthen America before it is too late. Like the previous two examples from the UK and India sub-corpora, the fellow commenters and readers were included by the commenter while advising the things that should be done.

Comparisons of the occurrences of *engagement markers*

From Figure 7.10, we extracted some notable characteristics of using *engagement markers* in the comments of the India, UK, and USA commenters.

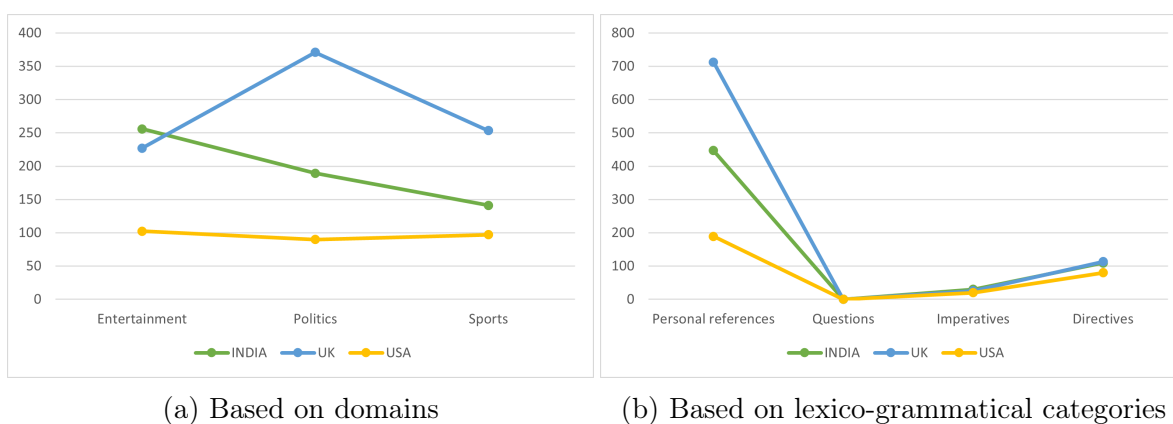


Figure 7.10: Comparisons of the occurrences (normalized) of *engagement markers*

- From Figure 7.10a, we can see that the usage patterns obtained with all the countries across the three chosen domains were different. The UK commenters employed *engagement markers* most in the politics domain, whereas the commenters from India and the USA used them most in the entertainment domain. Notably, the USA commenters used *engagement markers* almost in the same frequency across all the domains (with a slight variation), whereas the use of this

marker across different domains by the UK and India commenters was significantly different. One plausible reason for this could be that the USA commenters of the entertainment, politics, and sports domains were keener to show solidarity with their readers and commenters and to introduce their thoughts in the conversations. Also, it is evident that the UK commenters applied *engagement markers* extensively, followed by the India and USA commenters which suggests the tendency of the UK commenters to engage the readers and other commenters in the arguments.

- Analysing our corpus based on the lexico-grammatical categories of *engagement markers*, we conclude from the Figure 7.10b that the usage patterns were identical, more specifically, the pattern was generated following the usage of the lexico-grammatical categories of *engagement marker* in the following order: personal references, directives, and imperatives. Personal references were used the most by all the commenters of the three chosen countries; more specifically, they were extensively used by the UK commenters in comparison to India and USA commenters which highlights the inclusive nature of the commenters where they include the readers and the other commenters in the conversation by expressing their thoughts and concerns. Notably, questions were not used by any of the commenters, irrespective of domains and political ideology, which suggests that the commenters were not likely to raise questions to increase the curiosity of the readers and other commenters in the form of questions. Imperatives were also used on a few occasions by commenters from all the nations.

7.5.5 Interactional: Self-mentions

Self-mentions is one of the crucial metadiscourse markers used by the authors for self-projection, claiming their presence within the text and influencing the readers of that particular text. *Self-mentions* are studied enormously by some of the writers belonging to different genres, such as Hyland (2001), Dueñas (2007), Walková (2019), Albalat-Mascarell and Carrió-Pastor (2019) where they have studied self-mentions as a medium to proclaim authorial stance and authorial self. In our study, *self-mentions* were used in the following ways: 1) to reveal the author’s involvement (in our case, commenter’s) in a particular text (in our case, comments) using first-person pronouns and self-citations. 2) to show personal engagement and self-representation with the audience while commenting on a topic or issue. We found that in our corpus, *self-mentions* were used less by the commenters compared to the other interactional metadiscourse markers, except for the India sub-corpora. While comparing with other studies, we observed that, unlike our study, in the study of Mur-Dueñas (2011), *self-mentions* were used significantly by the writers. Table 7.9 illustrates the normalized frequencies of *self-mentions* found in India, the UK, and the USA sub-corpora. In this section, we also describe the frequency of the lexico-grammatical categories of *self-mentions*: personal reference and self-citation.

Table 7.9: Normalized frequencies of *self-mentions* found in India, the UK and the USA

		Entertainment			Politics			Sports		
		LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
INDIA	Personal reference	11.09	53.08	64.17	38.93	123.21	162.14	82.35	84.11	166.46
	Self-citation	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total	11.09	53.08	64.17	38.93	123.21	162.14	82.35	84.11	166.46
UK	Personal reference	1.67	11.55	13.21	3.50	102.47	105.96	1.00	18.57	19.57
	Self-citation	0.00	0.00	0.00	0.00	0.94	0.94	1.00	0.00	1.00
	Total	1.67	11.55	13.21	3.50	103.40	106.90	1.99	18.57	20.57
USA	Personal reference	48.29	43.19	91.48	48.84	60.21	109.05	37.27	65.73	103.00
	Self-citation	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total	48.29	43.19	91.48	48.84	60.21	109.05	37.27	65.73	103.00

The following occurrences from the lexico-grammatical categories of *self-mentions* were used by India, the UK, and the USA commenters irrespective of domains and

political ideology: personal reference (*we, our*) and self-citation (*my, author*).

India: From Table 7.9, we can see the normalized frequencies of *self-mentions* used in the India sub-corpora irrespective of domain and political ideology. The general use of the lexico-grammatical categories of *self-mentions* was seen more in the sports domain, followed by the politics and entertainment domain. In the sports domain, regarding political ideology, the left-wing and right-wing commenters used personal references almost the same, whereas self-citations were never used by them a single time which suggests the fact that the commenters of both the left-wing and right-wing were disinterested to show their authorial identity. In the politics domain, there was a significant difference in using personal references among the left-wing and right-wing commenters, where the right-wing commenters used them almost 3 times more than the left-wing. In the entertainment domain also, the right-wing commenters used personal references more than 4 times than the left-wing ones. Interestingly, the India commenters never used self-citations, irrespective of domains and political ideology which upholds the indifference on the part of the commenters to show their authorial selves.

The UK: The general use of the categories of *self-mentions* was seen remarkably more in the politics domain, followed by the sports and entertainment domain. In the politics domain, the left-wing commenters used personal references significantly less (3.50 per 10,000 words) compared to the right-wing commenters (102.47 per 10,000 words). This difference indicates the tendency of the right-wing commenters of the politics domain to include the readers' presence in the argument. Also, we noticed that self-citations were not used a single time by the left-wing commenters, whereas it was used significantly less by the right-wing commenters (0.94 per 10,000 words). In the sports domain and entertainment domains, the right-wing commenters used personal references comparatively more than the left-wing. Unlike the politics and sports domain, the left-wing and right-wing commenters of the entertainment domain did not use self-citation for once which shows that the commenters of the entertainment

domain were not so interested in projecting authorial selves in the argument.

The USA: The USA commenters used the lexico-grammatical categories of *self-mentions* more frequently in the politics domain, followed by the sports and entertainment domain. Regarding political ideology, in the politics and sports domains, the commenters of the right-wing sub-corpora used personal references more than the left-wing. Also, irrespective of political ideology, in the politics and sports domain, the commenters did not use any self-citations. In the entertainment domain, the left-wing commenters used personal references a bit more. However, self-citations were never used by both the left-wing and right-wing commenters. The absence of self-citation in the USA sub-corpora, irrespective of domain and political ideology, points out the fact that the USA commenters were more into highlighting collective identity rather than projecting their authorial selves. Also, the commenters could think of the idea of self-projection as very subjective, which is not appropriate for communication and interaction.

Occurrences in Lexico-Grammatical categories

Table 7.10 portrays the quantitative analysis of the occurrences from each lexico-grammatical category of *self-mentions*. It includes the frequencies from the India, UK and USA sub-corpora. As the number of words was less, we thought to mention all the occurrences from the three nations.

Table 7.10: Top-5 (normalised frequency per 10,000 words) words of lexico-grammatical categories of *self-mentions*

	INDIA		UK		USA	
Personal ref.	we	250.79	we	106.98	we	228.10
	our	34.42	our	19.92	our	20.69
Self-citation			my	0.99		
			author	0.67		

Personal reference: The first lexico-grammatical category of *self-mentions* is personal reference. In this category, mainly exclusive pronouns (such as ‘we’ and ‘our’) are found that exclude the involvement of the listener and emphasize the involvement

of the writer as well as the readers. Only two personal references occurred in the India, UK, and USA sub-corpora— ‘we’ and ‘our’. Personal references were used more frequently in the India sub-corpora, followed by the USA and UK sub-corpora. In the India and USA sub-corpora, the first-person plural pronoun ‘we’ were used significantly, whereas, in the UK sub-corpora, they were used comparatively less. Similarly, another first-person plural pronoun ‘our’ were used comparatively more in the India sub-corpora (34.42 per 10,000 words), while they were used a bit less in the USA and UK sub-corpora (20.69 and 19.92 per 10,000 words, respectively). Here, we show some examples of the exclusive pronoun ‘we’ from our corpus.

Example 7.5.56 *“Er not the case. We currently have an unemployment rate of just 4.3 at present. Income is the same on average for immigrants v UK born in both males and females. We also have one of the worst fertility rates in the developed world. This means that we will need to do something to encourage immigration to the UK or, in time, people’s pensions won’t get paid.” (source:UK-RW-POL:1)*

In this example from the UK sub-corpora, the commenter argued about the effects of Brexit. He talked about some issues that the UK was facing at the time of Brexit. Here, the first-person plural pronoun ‘we’ was used by the commenter while replying to the other commenter. It is also an exclusive pronoun that was used by the commenter to mean the commenter himself and the readers and other commenters but excludes the commenter or the listener with whom the commenter was replying.

Example 7.5.57 *“Every artistic people has some personality disorder (ego) which they do not expose outside. So we viewer shouldn’t take it as negative but appreciate him for giving us a smiling face.” (source:IND-RW-ENT)*

While answering another commenter in a conversation regarding the famous Indian comedian Kapil Sharma, the commenter of the India sub-corpora used the exclusive pronoun ‘we’. It was used by the commenter to mean the commenter himself and the

readers and other commenters but excludes the listener or the commenter with whom he was having an argument. He included the readers and other commenters in the comment to state that in this tough time for Kapil Sharma, everyone should support him instead of criticizing him.

Example 7.5.58 “*We already had Republicans trying to tell us how to "fix" our state. Spoiler alert, they didn't. We're doing fine now. Not perfect, no, but better than KS. Or any of the other states having issues with teachers strikes right now: OK, KY, WV, I could go on..*” (*source:USA-LW-ENT*)

In this example from the USA sub-corpora, the commenter argued that earlier, the citizens of the USA had Republicans who tried to tell them how to run the state successfully, though they had not done anything to make the nation better. Here, the commenter used the exclusive pronoun ‘we’ while arguing with the other commenter. He deliberately excluded the particular commenter with whom he was having an argument and included the readers and other commenters who were reading this comment.

Self-citations: The second lexico-grammatical category of *self-mentions* is self-citation, where the author refers to himself or his/her work. In our study, *self-citations* were used to mean something that belongs to the author or the speaker or, in our case, the commenter. We found only two self-citations from our corpus – ‘my’ and ‘author’. These occurrences were found only in the UK sub-corpora as the India and USA commenters never used them a single time. Here, we mention some examples of self-citations from the UK sub-corpora.

Example 7.5.59 “*I'm so frustrated - my wife cancelled my kid's swimming lesson at the last minute so I had to stay at home and watch the match instead. As for the football, I blame Cahill and Fabregas.*” (*source:UK-LW-ENT*)

In this example, the commenter showed his frustration with the present situation where he had to watch a football match while staying at home instead of joining his

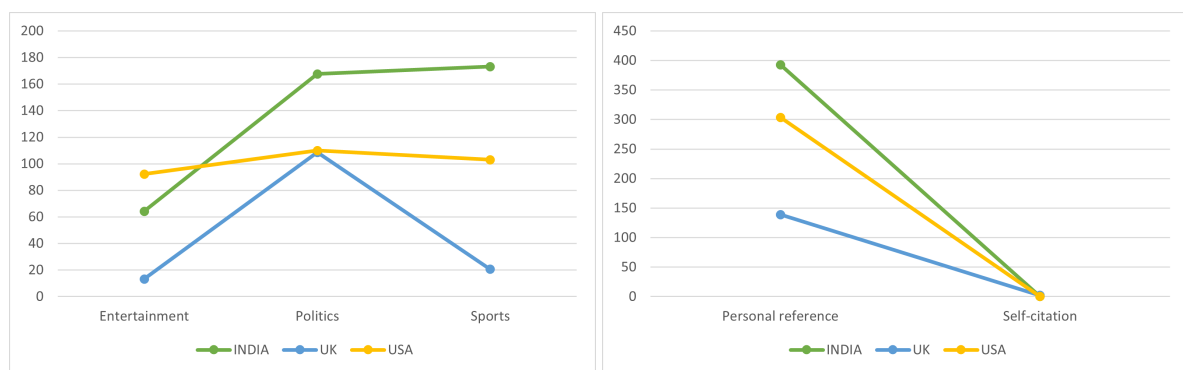
kids for swimming as his wife canceled the swimming lesson. In this comment, the first-person single pronoun ‘my’ was used to mean something or someone that belongs to the commenter himself. ‘My’ is also a possessive pronoun and possessive adjective as well, as it is used to replace the possessive form of the noun (in this case, the commenter himself). It was used by the commenter instead of using the personal pronoun ‘I’ to signify possession, i.e., his wife and his kids.

Example 7.5.60 “*He may well be ‘a silly little fellow’ but he’s not an author of odious repugnant posts is he David.*” (source:UK-RW-POL:1)

In this example from the UK sub-corpora, the commenter referred to a fellow commenter, who, according to the commenter, was not the writer of some repulsive posts. Here, he was defending a particular fellow commenter from his other fellow commenters. The noun ‘author’ was used by the commenter to mean the writer of the comment. The commenter used it to project the authorial self of the writer (in our case, the particular commenter).

Comparisons of the occurrences of *self-mentions*

From Figure 7.11a, we extracted some notable characteristics of using *self-mentions* in the comments by the commenters of India, the UK, and the USA.



(a) Based on domains

(b) Based on lexico-grammatical categories

Figure 7.11: Comparisons of the occurrences (normalized) of *self-mentions*

- From Figure 7.11a, we can see that the usage pattern obtained with all the countries across the three chosen domains was almost similar, except for India. The use of *self-mentions* was less compared to the other metadiscourse markers which shows the indifferent nature of the commenter to mention his or her presence in the conversation. In the politics domain, *self-mentions* were used the most among all the domains irrespective of nations except India, where it was used a bit more in the sports domain. The USA commenters used *self-mentions* almost the same for all the domains, with a slight difference. The use of *self-mentions* was almost the same by the UK and USA commenters in the politics domain, however, the India commenters used them comparatively more than the UK and the USA which points out the possessive nature of the commenters who used exclusive and possessive pronouns in their comments to include the readers and other commenters in the conversation.
- Analysing our corpus based on the categories of *self-mentions*, we conclude from the Figure 7.11b that the usage patterns of different categories for three chosen nations were identical, more specifically the pattern generated following the usage of the *self-mentions* in the following order: personal reference and self-citation. However, the usage was different across different countries, more specifically, the India and USA commenters used them more frequently than the UK commenters. Personal references were used the most by all the commenters of the three chosen countries; more specifically, they were extensively used by the India commenters, followed by the USA and UK commenters. Notably, self-citations were used a way behind the personal-references, and surprisingly they were used only by the UK commenters. This suggests the tendency of the UK commenters not to show their authorial self while commenting on digital news.

7.5.6 Interactive: Transition markers

Transition markers are one of the prime categories of interactive metadiscourse markers used by writers for better comprehension by providing a link between propositions, ideas, sentences, and paragraphs of a text. The writers use them to connect ideas and show the relationship between ideas. These markers show the additive, contrastive and consecutive relation between ideas of a text. Previous studies such as Cao and Hu (2014), Hyland (2005a), and Hyland and Jiang (2018) have explored *transition markers* to show the relationship between ideas of a text. In our study, *transition markers* were used by the India, UK, and USA commenters, irrespective of domains, and political ideology in the following ways: 1) to provide a link between the ideas and comments. 2) to build a well-organized text to help the commenters to interpret the links between ideas. This section describes the in-depth analysis of *transition markers* and its lexico-grammatical categories. From our corpus, regarding the usage, we found that among the interactive metadiscourse categories, *transition markers* have the highest frequency. In this case, our study falls in similar line with the study of Hyland (2005a). The following lexico-grammatical categories of *transition markers* were used by the India, UK, and USA commenters: additive (*and, as well, so, as well as, also, further, likewise, furthermore, moreover, in addition, similarly*), contrastive (*but, instead, though, yet, rather, however, still, otherwise, nevertheless, by contrast, alternatively, on the other hand*) and consecutive (*so, then, so that, as a result, thus, as such, following, next, for any reason, hence, consequently*).

India: From Table 7.11, it is evident that the India commenters of the politics domain overall used *transition markers* more, followed by the sports and entertainment domains, which suggests that the commenters of the politics domain were more ardent in letting know their readers and other commenters the underlying connections between ideas and propositions. In the politics domain, the overall use of the lexico-grammatical categories of *transition markers* was seen more in the left-wing sub-corpora. Additive and consecutive markers were used by the left-wing commenters comparatively more

Table 7.11: Normalized frequencies of *transition markers* found in India, the UK, and the USA

		Entertainment			Politics			Sports		
		LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
INDIA	Additive	141.35	108.81	250.16	163.30	96.03	259.33	171.38	84.11	255.49
	Contrastive	49.89	61.04	110.93	70.29	72.48	142.77	55.64	65.42	121.06
	Consecutive	11.09	47.77	58.86	38.93	32.61	71.55	35.61	23.36	58.98
	Total	202.33	217.62	419.95	272.52	201.12	473.64	262.63	172.90	435.53
UK	Additive	163.33	100.08	263.41	132.91	105.74	238.65	172.36	114.09	286.45
	Contrastive	113.33	111.62	224.96	68.79	81.88	150.67	104.61	66.33	170.94
	Consecutive	35.00	19.25	54.25	55.96	50.53	106.49	46.83	45.10	91.93
	Total	311.67	230.95	542.61	257.67	238.15	495.82	323.80	225.52	549.33
USA	Additive	110.60	82.27	192.87	113.02	110.15	223.18	138.91	114.22	253.14
	Contrastive	46.73	78.16	124.89	60.77	84.29	145.06	88.09	79.74	167.83
	Consecutive	28.82	26.74	55.56	35.78	31.22	67.00	38.96	42.03	80.99
	Total	186.15	187.17	373.32	209.58	225.66	435.24	265.97	235.99	501.96

than the right-wing. However, the use of contrastive markers was almost the same by the left-wing and right-wing commenters, with a slight difference. In the sports domain, the overall use of the categories was substantially more in the left-wing compared to the right-wing. Additive markers were used by the left-wing commenters significantly more than the right-wing, whereas contrastive markers were used comparatively more by the right-wing. The use of consecutive markers was on the lower side for both the left-wing and right-wing sub-corpora; however, the left-wing commenters used them a bit more. In the entertainment domain, the overall use of *transition markers* was more in the right-wing sub-corpora. In the use of additive markers, the left-wing commenters used comparatively more than the right-wing; contrastive markers were seen to be used more in the right-wing sub-corpora than the left-wing. A difference was noticed in the use of consecutive markers, where the right-wing commenters used them almost 4 times more than the left-wing (47.77 and 11.09 per 10,000 words). One could think as a possible explanation that the right-wing commenters of the entertainment domain were keener on predicting the consequence of a situation and wanted to inform their readers about them beforehand.

The UK: The UK commenters used the lexico-grammatical categories of *transition markers* more in the sports domain, followed by the entertainment and politics domains. In all three domains, the lexico-grammatical categories of *transition mark-*

ers were used more in the left-wing sub-corpora. In the sports domain, the left-wing commenters used additive and contrastive markers comparatively more than the right-wing, whereas, in the use of consecutive markers, the use was almost identical in both the left-wing and right-wing sub-corpora, with a slight difference. In the entertainment domain, additive and consecutive markers were used comparatively more in the left-wing compared to the right-wing. However, the use of contrastive markers was almost the same in both the left-wing and right-wing sub-corpora which suggests that the commenters of both the left-wing and right-wing were more eager to show the similar and contrastive relations between the main clauses from the comments. In the politics domain, additive markers were used comparatively more in the left-wing sub-corpora; on the contrary, contrastive markers were used more by the right-wing commenters. However, consecutive markers were used slightly more by the left-wing commenters. This indicates the tendency of the commenters of the politics domain to make the readers and other commenters understand the transition within the comments.

The USA: The quantitative analysis of the lexico-grammatical categories of *transition markers* reveals that in general use, the commenters of the sports domain used the categories more frequently, followed by the politics and entertainment domains. In the sports domain, the left-wing commenters used the categories a bit more than the right-wing. They used additive and contrastive markers more frequently than the right-wing ones. The use of consecutive markers was on the lower side in both the sub-corpora; however, it was used a bit more by the right-wing commenters which indicates the tendency of the commenters of the sports domain to suggest the consequences beforehand to the readers and other commenters. In the politics domain, the right-wing commenters used the categories overall more compared to the left-wing. Additive markers were used almost the same in both the sub-corpora. In the use of contrastive markers, the right-wing commenters used the categories comparatively more (84.29 per 10,000 words) than the left-wing commenters (60.77 per 10,000 words). Consecutive markers were used a bit more in the left-wing than in the right-wing. In

the entertainment domain, we found that irrespective of political ideology, additive and consecutive markers were used more frequently by the left-wing commenters. In contrast, contrastive markers were used comparatively more by right-wing commenters.

Occurrences in Lexico-Grammatical categories

This section will discuss the in-depth analysis of the occurrences from each lexico-grammatical category of *transition markers*. Table 7.12 highlights the top five occurrences of the lexico-grammatical categories of *transition markers*.

Table 7.12: Top-5 (normalized frequency per 10,000 words) words of lexico-grammatical categories of *transition markers*

	INDIA		UK		USA	
Additive	and	646.90	and	679.24	and	560.02
	also	77.62	also	39.23	also	35.36
	so	21.68	so	31.54	so	34.94
	as well	12.65	as well	24.32	as well	13.58
	further	2.16	therefore	6.10	as well as	7.92
Contrastive	but	269.52	but	351.80	but	307.84
	still	29.20	though	60.80	still	36.04
	instead	21.69	still	49.21	though	27.61
	otherwise	15.55	instead	25.07	instead	19.86
	rather	12.50	yet	22.54	however	16.38
Consecutive	then	90.47	so	121.81	so	99.38
	so	55.32	then	114.32	then	79.34
	so that	17.85	so that	3.10	so that	4.75
	hence	11.72	as a result	1.93	thus	4.65
	thus	6.39	hence	1.40	as a result	3.74

Additive: The first lexico-grammatical category of *transition markers* is additive, which provides additive relation between main clauses (Hyland and Jiang (2018)). Additive markers are usually applied by the writers through adverbs and conjunctions. We found that in the three sub-corpora, the four most frequently used additive markers were ‘and’, ‘also’, ‘so’, and ‘as well’. The additive markers ‘and’ and ‘as well’ were most frequently used by the UK commenters. Another additive marker, ‘also’, was used most frequently by the India commenters, followed by the UK and USA commenters. In the use of ‘so’, we found that the USA commenters used it more frequently, followed by the commenters of the UK and India. Apart from these four additive markers, the India, UK, and USA commenters frequently used ‘further’, ‘therefore’, and ‘as well’.

as' respectively in their comments. Here, we discuss some examples of 'and' from our corpus to understand the function of additive markers.

Example 7.5.61 *“You’re spot on mate, that’s why we’ve voted to leave those EU features behind us and cultivate a more open, world friendly free trading nation with other countries who have the same values.” (source:UK-RW-POL:1)*

Example 7.5.62 *“Mary is a great boxer in the 48 kg category for women. The 48 kg category for women was abolished in the Olympics and Mary was forced to move up to the 51 kg category. In this category she was easily outmanoeuvred by the likes of Michaela Adams who won the Olympic gold in 2012 and 2016.” (source:IND-LW-SPR)*

Example 7.5.63 *“I’m happy the Russians won. Putin is already livid because of the sanctions, no need to make him any madder (and silver is really really good for Germany).” (source:USA-LW-SPR)*

In the examples mentioned above from the three sub-corpora, the additive marker 'and' was used in a two-fold way: 1) to join two propositions and 2) to introduce additional information. In the first example from the UK sub-corpora, the commenter used the conjunction 'and' to join two propositions while describing his fellow commenter about the benefits of Brexit that they will have in trade and commerce. In the second example from the India sub-corpora, the commenter used 'and' to join two propositions where he explained the reason for Mary Kom's (an Indian boxer) loss at the Tokyo Olympics. In the third example from the USA sub-corpora, the commenter used 'and' to include additional information. Here, the commenter explained his thought regarding Germany's getting silver, and he added additional information in the brackets. In all three above examples, the commenters from the UK, India, and the USA sub-corpora, used the conjunction 'add' to continue the flow of the conversation.

Contrastive: The second lexico-grammatical category of *transition markers* is contrastive marker, which is used by writers to project contrastive relations between ideas,

propositions, and main clauses. Contrastive markers are realized by the writers through conjunctions and adverbs expressing contrast. From the India, UK, and USA sub-corpora, three contrastive markers ‘but’, ‘still’, and ‘instead’, were found to be common. The conjunction ‘but’ was substantially used in all three sub-corpora (UK: 351.80 per 10,000 words, USA: 307.84 per 10,000 words, and India: 269.52 per 10,000 words). The adverb ‘still’ was used more frequently in the UK sub-corpora, followed by the USA and India sub-corpora. Another adverb, ‘instead’ was used most frequently by the UK commenters, followed by the India and USA commenters. So, in all three cases, the use of contrastive markers was more in the UK sub-corpora. The other two contrastive markers used in the India sub-corpora were ‘otherwise’ and ‘rather’; in the UK sub-corpora, ‘though’ and ‘yet’; and in the USA sub-corpora, ‘though’ and ‘however’. Here, we cite some examples of the conjunction ‘but’ from the three sub-corpora.

Example 7.5.64 *“Another one ‘flaunting’ her incredible, sensational figure. Yes she does have a good figure but so do millions of other girls. The vanity of this one (whoever she is) is just horrible.” (source:UK-RW-ENT)*

Example 7.5.65 *“The first assassinate of Free India was a Saffron, not a Hindu. The people Killed Indira Gandhi were not Sikhs but Terrorists. Those who kill and bombs do not belong to any religion, they just misuse religion. Terrorists are cowards. Extremist do not follow any religion. They interfere in other people religion to create hatred and blood bath.” (source:IND-LW-POL)*

Example 7.5.66 *“Interestingly, PETA came out furious when some guy killed a shark in Florida. But they were dead silent (pardon the pun) when one of their own tried to slaughter human beings at YouTube.” (source:USA-RW-ENT)*

The commenters used ‘but’ in the above-mentioned examples to contrast one positive and one negative proposition. In the first example from the UK sub-corpora, the commenter used the conjunction ‘but’ to contrast other girls who also had great

physiques, just like Amber Turner (a model and TV personality). The commenter criticized the show off of the model and used ‘but’ to contrast one positive and one negative aspect. In the second example from the India sub-corpora, the commenter mentioned that the people who killed Indira Gandhi (a famous Indian politician and the then Prime Minister of India) were not Sikhs (the followers of the Sikh religion) as they were only terrorists (people who use unlawful violence against civilians). The commenter contrasts his remark by using the conjunction ‘but’ to contrast one positive (in this example, the Sikhs) and one negative (in this example, terrorist) proposition. In the third example from the USA sub-corpora, the commenter used ‘but’ to include a piece of information suggesting one contrastive aspect. The commenter described an incident where an organization named PETA (People for the Ethical Treatment of Animals) was extremely angry after knowing the death of a shark. He further mentioned that contrary to this situation, the same organization remained silent when they were doing the same mishap with human beings on YouTube. Here, ‘but’ contrasts one positive and negative aspect of PETA.

Consecutive: The third lexico-grammatical category of *transition markers* is consecutive marker, which writers use to express cause and effect relation. Also, these markers project consequential relations between main clauses (Hyland and Jiang (2018)). From our corpus, we found some common consecutive markers that occurred in the UK, India, and the USA sub-corpora: ‘then’, ‘so’, and ‘so that’. Consecutive markers are employed through adverbs by the writers. The adverb ‘then’ was used more frequently in the UK sub-corpora followed by the India and USA sub-corpora. Another adverb, ‘so’, was used more in the UK sub-corpora, followed by the USA and India sub-corpora. ‘So that’ was used comparatively less in the UK and USA sub-corpora; however, they were used almost 4 times more in the India sub-corpora. The other consecutive markers that were used in the corpus were ‘hence’, ‘thus’, and ‘as a result’. Here we discuss some examples of the adverb ‘then’ from our corpus.

Example 7.5.67 *“Typical of our matches this year. Should have won easily, but small*

offsides, bad finishing and a blinder from Joe Hart. As soon as Hernandez came on, I told a friend that he would score. Willian missed very closely and then seemed afraid to shoot, and Pedro decided he should hit a few into the crowd. At least, as opposed to Arsenal, the stadium was full.” (source:UK-LW-SPR)

Example 7.5.68 *“nothing adds up: They were at the wedding, ok good. Sri Devi stays back in UAE while husband and daughter returns to Mumbai. Why did she stay back w her husband. Then husband returns to ‘surprise’ her at 5.30pm. Chatted for 15 mins. Then bath time and then dinner outing. She was asleep during day time and alone. Would she be still alive. Do u think she took pills. The mystery is why she wanted to stay alone in a strange country. One can still rest at home in a familiar place. I am so confused.” (source:IND-LW-ENT)*

Example 7.5.69 *“The Republicans are the party that says government doesn’t work and then they get elected and prove it.” (source:USA-LW-POL)*

In the above-mentioned three examples from the three sub-corpora, the consecutive marker ‘then’ was used to mean something next, after a previous action. In the first example from the UK sub-corpora, the commenter described football matches where the commenter talked about a football player who missed very closely and afterward was scared to shoot further. In the second example from the India sub-corpora, the commenter described the incidents that took place one after another on the day of Sridevi’s (a famous Indian film actress) death. Here, the commenter used the adverb ‘then’ to mean something after a previous action. In the third example from the USA sub-corpora, the commenter criticized the Republicans (conservative and right-leaning) and stated that they prove it right when they themselves criticize other parties for not doing any meaningful work, and after getting elected do the same thing. Likewise the previous two examples, the commenter used ‘then’ to mean next or after a previous action.

Comparisons of the occurrences of *transition-markers*

Figure 7.12a depicts notable features of using *transition markers* by the India, UK, and USA commenters.

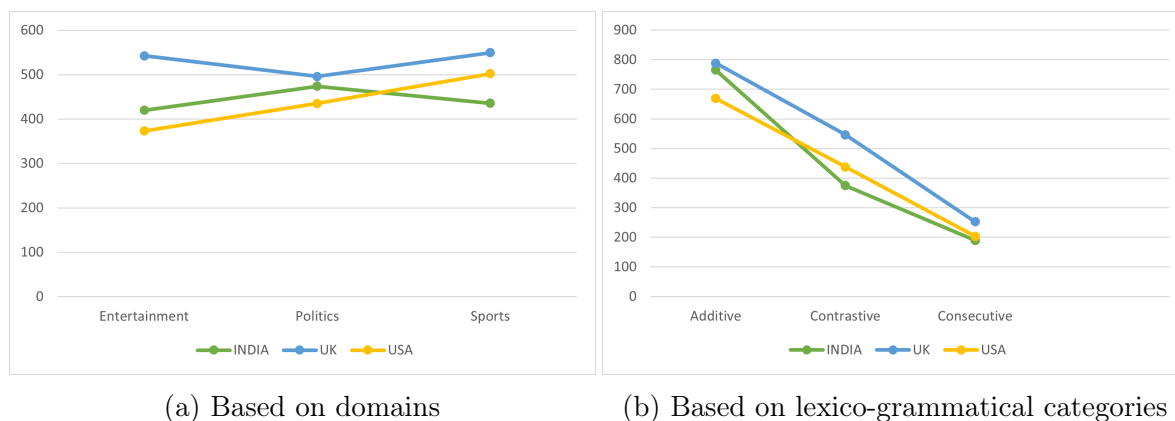


Figure 7.12: Comparisons of the occurrences (normalized) of *transition markers*

- From figure 7.12a, it is evident that the usage patterns of *transition markers* were different for the three domains. The India commenters used them most in the politics domain, whereas the UK commenters used them less (in the third place), and the USA commenters used them in the second place. Instead, in the sports domain, the UK and USA commenters used *transition markers* most frequently. Also, we noticed that in the entertainment domain of the India and USA sub-corpora, *transition markers* were used least frequently by the commenters. It indicates that the India and USA commenters of the entertainment domain were not so eager to demonstrate the transitions in the comments to their readers and other commenters. Perhaps, these commenters were more into encouraging their readers and other commenters without marking the transitions to interpret the comments by themselves.
- Figure 7.12b depicts that the usage pattern for the three lexico-grammatical categories of *transition markers* was the same, irrespective of countries. In the UK, India, and the USA sub-corpora, the additive markers were used most frequently;

contrastive markers were used second most frequently, and consecutive markers were used least frequently by the commenters of the three respective countries; however, differences were noticed in the use of these lexico-grammatical categories by the UK, India, and the USA commenters. It shows the tendency of the commenters of the UK, India, and the USA to show additive and contrastive relations between the comments; however, the less use of consecutive markers points out the fact that these commenters were not so eager to highlight the consequences in the comments.

7.5.7 Interactive: Frame markers

The second category of *interactive* metadiscourse marker is the *frame markers*. The writers use *frame markers* to indicate different stages in the text. According to Hyland and Jiang (2020, p. 5), “frame markers are references to rhetorical elements of text structure, used to sequence, label text stages, to announce discourse goals and to indicate topic shifts.” These markers indicate different functions in the text, for example, sequencing, shifting topics, etc. In our study, the commenters used *frame markers* in the following ways: 1) to indicate different functions of comment, including sequencing, shifting topic, and announcing ideas, etc. 2) to present the readers and other commenters a panoramic view of the comments. The lexico-grammatical categories of *frame markers* are sequencers, topicalizers, discourse labels, and announcers. In this section, we discuss the in-depth analysis of each of the lexico-grammatical categories of *frame markers*, irrespective of country, domain, and political ideology. In our study, *frame markers* are the second highest frequent category among the *interactive markers*. The following lexico-grammatical categories of *frame markers* were found in the UK, India, and USA sub-corpora: sequencers (*first, second, next, finally, third, to begin with*), topicalizers (*regarding, with respect to, in terms of, with regard to*), discourse labels (*overall, in short, conclude, in all, briefly*), announcers (*will, shown, seek, aim to*). Table 7.13 portrays the normalized frequencies of the *frame markers* found in the

UK, India, and the USA sub-corpora.

India: Table 7.13 depicts that the lexico-grammatical categories of *frame markers* were used more generally in the sports domain, followed by the entertainment and politics domains. In the sports domain, the categories were used overall more (almost 2 times) by the right-wing commenters than the left-wing. Sequencers and announcers were used comparatively more in the right-wing sub-corpora. Topicalizers were not used for a single time in any of the sub-corpora of the sports domain which shows that the commenters were aloof in signaling the readers about the topic shifts in arguments. The left-wing commenters never used discourse labels once in their comments; however, they were used by the right-wing commenters for 18.69 per 10,000 words. In the entertainment domain, we noticed an interesting fact in the use of the categories by the left-wing commenters, where they did not use sequencers, topicalizers, and discourse labels for a single time. The right-wing commenters used sequencers for 13.27 per 10,000 words and topicalizers for 2.65 per 10,000 words; however, they never used discourse labels for once. The left-wing commenters used announcers significantly more than the right-wing. Overall, the right-wing commenters of the politics domain used the categories of *frame markers* significantly more. They used sequencers almost 3 times more than the left-wing. Topicalizers and discourse labels were never used in the left-wing sub-corpora; however, discourse labels were used very less by the right-wing commenters (1.81 per 10,000 words). Announcers were almost 4 times more in the right-wing sub-corpora than the left-wing that indicates the right-wing commenters were more into letting the readers and other commenters know the direction of the argument.

The UK: The quantitative analysis of the lexico-grammatical categories of *frame markers* by the UK commenters indicates that the commenters of the politics domain overall used *frame markers* more, followed by the sports and entertainment domains. In the politics domain, the right-wing commenters overall used *frame markers* comparatively more than the left-wing. The left-wing commenters used sequencers more

than the right-wing. Topicalizers were used comparatively more in the right-wing sub-corpora. Discourse labels were never used by the left-wing commenters; on the other hand, the right-wing commenters used them very little (0.47 per 10,000 words). In the sports domain, sequencers and topicalizers were used 16.94 and 1.99 per 10,000 words by the left-wing commenters, respectively, whereas there was not a single use of them in the right-wing sub-corpora. Contrarily, discourse labels and announcers were used comparatively more in the right-wing sub-corpora. In the entertainment domain, sequencers and announcers were used slightly more in the left-wing sub-corpora. Topicalizers and discourse labels were never used for a single time in both the sub-corpora. This shows the commenters' indifference in letting the readers know about the topic shifts and drawing the attention of the readers and other comments in an argument.

The USA: The commenters of the sports domain used *frame markers* more, followed by the entertainment and politics domain. In the sports domain, the right-wing commenters used all the lexico-grammatical categories of *frame markers* more, except for sequencers. The left-wing commenters used sequencers comparatively more than the right-wing. They did not use topicalizers and discourse labels a single time; however, both categories were used very less by the right-wing commenters. In the entertainment domain, the use of sequencers, topicalizers, and discourse labels was on the lower side by both left-wing and right-wing commenters. Sequencers were not used for a single time in the right-wing sub-corpora; however, they were used very less (1.56 per 10,000 words) in the left-wing sub-corpora. The use of topicalizers was on the lower side for both the left-wing and right-wing, though it was used a bit more by the right-wing commenters. The left-wing commenters did not use discourse labels a single time; on the contrary, the right-wing commenters used them for 2.06 per 10,000 words. Announcers were used comparatively more in the left-wing sub-corpora. In the politics domain, the right-wing commenters used the categories of *frame markers* more. Sequencers and announcers were used comparatively more by right-wing. However, we found that topicalizers and discourse labels were not used once by the left-wing

commenters, whereas they were used on the lower side by the right-wing. Announcers were used comparatively more by the right-wing which reveals the tendency of these commenters to make the readers and other commenters understand the direction of the argument.

Table 7.13: Normalized frequencies of *frame markers* found in India, the UK, and the USA

		Entertainment			Politics			Sports		
		LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
INDIA	Sequencers	0.00	13.27	13.27	6.49	18.12	24.61	8.90	14.02	22.92
	Topicalizers	0.00	2.65	2.65	0.00	0.00	0.00	0.00	0.00	0.00
	Discourse labels	0.00	0.00	0.00	0.00	1.81	1.81	0.00	18.69	18.69
	Announcers	72.06	42.46	114.52	19.47	74.29	93.75	40.06	65.42	105.48
	Total	72.06	58.39	130.45	25.95	94.22	120.17	48.97	98.13	147.10
UK	Sequencers	8.33	7.70	16.03	17.49	10.76	28.25	16.94	0.00	16.94
	Topicalizers	0.00	0.00	0.00	2.33	2.81	5.14	1.99	0.00	1.99
	Discourse labels	0.00	0.00	0.00	0.00	0.47	0.47	1.00	2.65	3.65
	Announcers	25.00	23.09	48.09	50.13	64.57	114.70	25.90	31.84	57.74
	Total	33.33	30.79	64.13	69.95	78.60	148.56	45.83	34.49	80.32
USA	Sequencers	1.56	0.00	1.56	5.68	6.24	11.92	13.55	8.62	22.17
	Topicalizers	1.56	2.06	3.61	0.00	0.89	0.89	0.00	3.23	3.23
	Discourse labels	0.00	2.06	2.06	0.00	1.34	1.34	0.00	1.08	1.08
	Announcers	38.16	22.62	60.79	15.33	26.31	41.65	35.58	80.82	116.39
	Total	41.28	26.74	68.02	21.01	34.79	55.80	49.13	93.75	142.88

Occurrences in Lexico-Grammatical categories

In this section, we discuss the in-depth analysis of the occurrences from each lexico-grammatical category of *frame markers*. Table 7.14 highlights the occurrences of the lexico-grammatical categories of *frame markers*.

Sequencers: The first lexico-grammatical category of *frame marker* is sequencer. The writers use sequencers to order parts of a text or an argument. According to Cao and Hu (2014), sequencers are used to order discourse-internal units. Following Hyland and Zou (2020), we observed that the sequencers that we found in our corpus were mainly numerical sequencers and temporal sequencers. The most commonly used sequencers in India, UK, and USA sub-corpora were ‘first’, ‘second’, and ‘finally’. ‘First’ was used more in the India sub-corpora, followed by the UK and the USA sub-corpora. ‘Second’ was used comparatively more in the UK sub-corpora, whereas it was used almost the same in the India and USA sub-corpora. ‘Finally’ was used

significantly less in all the three sub-corpora; however, it was used a bit more (1.08 per 10,000 words) by the India commenters than by the UK (1.00 per 10,000 words) and USA commenters (0.45 per 10,000 words). We found the following other occurrences from our corpus: *third*, *next*, and *to begin with*. However, we are mentioning here the examples of the numerical sequencers as those were common in the UK, India, and the USA sub-corpora.

Example 7.5.70 *“Homophobic hate crime is up, and the extreme right are emboldened by Trump, Brexit and other divisive ‘populist’ movements. LGBT people are firmly in their crosshairs as their first targets in their quest to oppress minorities and ‘restore order’.” (source:UK-LW-ENT:2)*

Example 7.5.71 *“First Weightlifters, then Shooters and now wrestlers.. You guys have really made the nation proud. It just goes on to show that if the infrastructure and facilities provided to sportspeople is improved, then they can bring laurels to the country even in the Olympics...” (source:IND-RW-SPR)*

Example 7.5.72 *“Trump doesn’t care about effective policy. He cares about propping up a self image as a tough guy. Just remember: to understand Trump, just realize that his first and last interest is always serving himself.” (source:USA-RW-POL:1)*

In the examples mentioned above from the UK, India, and USA sub-corpora, the commenter used the sequencer ‘first’ to refer to an initial step or event. In the first comment, the UK commenter criticized the actions of the extreme right party, and in order to describe and order the future actions of the right followers, the commenter used the sequencer ‘first.’ In the second comment by the India commenter, ‘first’ was used to mention an initial category of Indian players who made the country proud by winning in the Commonwealth Games. He used ‘first’ to indicate a sequence. In the third example from the USA sub-corpora, the commenter used ‘first’ to refer to the prime or initial interest of the then President of America, Donald Trump. By mentioning Trump’s

interest in sequence, the commenter actually criticized the President's actions. In the three above examples, the numerical sequencer 'first' was used by the commenters to mean something that comes at the initial position.

Topicalizers: Topicalizers are used by writers to indicate changes from one topic to another, particularly referring to a previous topic that has been introduced before. Also, they are used to introduce a new topic and to include additional information regarding what was stated before. We found that from the three sub-corpora (UK, India, and USA), the commenters, irrespective of nation, used topicalizers very less in their comments. We observed four occurrences of topicalizers ('*in terms of*', '*regarding*', '*with respect to*', and '*with regard to*') from our corpus, and among these four, the topicalizer 'in terms of' occurred in all the three sub-corpora. The topicalizer 'in terms of' was used a bit more in the India sub-corpora, followed by the UK and USA sub-corpora. 'Regarding' was never used in the India sub-corpora; however, it was used more frequently in the UK and USA sub-corpora. The other two topicalizers 'with respect to' and 'with regard to' were not used by the India and USA commenters for a single time. However, 'with respect to' and 'with regard to' were used by the UK commenters for 1.17 and 0.47 per 10,000 words. 'In terms of' was used in all three corpora, so here we cite some examples from our corpus.

Example 7.5.73 *"You must be talking about 'huge' contracts between particular economic operators in Germany or Poland, and a buyer (which might be private or public sector) in whichever Commonwealth country. In terms of state-to-state trade agreements, however, neither Germany nor Poland can negotiate or conclude trade deals with sovereign counterparties. The European Commission has an 'exclusive competence' to negotiate and conclude such deals on behalf of the 28 EU member states."*
(source:UK-RW-POL:2)

Example 7.5.74 *"... I don't think that people can digest and enjoy stereotype misdemeanor from this arrogant guy. Except for initial days of "Laughter challenges"*

episodes he has always been on his constant downfall in terms of so-called comedy. You cannot stop the inevitable.” (source:IND-RW-ENT)

Example 7.5.75 *“At least the government was not involved. In "happy" Sweden she probably would be arrested and sent to a rehabilitaion center. I am exaggerating of course but we all forget how good we have it here in terms of expression.” (source:USA-LW-ENT)*

In the examples mentioned above from the three sub-corpora, the respective commenters used the topicalizer ‘in terms of’ to mean with regard to one particular aspect. The commenter here discussed trade and commerce after Brexit, especially the trade deals between the EU and the Commonwealth countries. After that, While describing the state-to-state trade deals between the EU countries, the commenter used the topicalizer ‘in terms of’ to shift from one topic to another or to introduce a new topic. In the second example from the India sub-corpora, the commenter used ‘in terms of’ to mean something concerning. The commenter criticized the Indian comedian Kapil Sharma for his arrogance and behaviour. Next, he praised Kapil Sharma for his work in the initial days of his career and compared it to his present downfall. The topicalizer ‘in terms of’ was used to introduce additional information here. In the third example from the USA sub-corpora, the commenter talked about Laura Ingraham (American Conservative television host), where the commenter used the topicalizer ‘in terms of’ to mean with regard to something. The commenter here compared the treatment that Laura Ingraham would have gotten for her actions if she was in Sweden and then compared it to the situation in the United States. The commenter used ‘in terms of’ to include additional information.

Discourse labels: The third lexico-grammatical category of *frame markers* is discourse labels which are used by the writers to label discourse stages. Usually, these markers help the writers to mark a more formal way of summarizing an idea. From our corpus, we found five discourse labels; however, we could not find any common occurrence among the three sub-corpora. We found three occurrences that were common

in two sub-corpora: 1) ‘In all’ were found in the India and USA sub-corpora, where in the USA sub-corpora, it was used significantly less. 2) ‘Overall’ were found in the India sub-corpora almost 4 times more than the UK sub-corpora. 3) ‘Conclude’ were used more in the India sub-corpora than in the USA. 4) ‘In short’ were used in the UK sub-corpora more than the USA sub-corpora. All the occurrences were on the lower side in the three sub-corpora. As there was no common occurrence among the three sub-corpora, we discuss some examples of ‘in short’, which is common in the UK and USA sub-corpora.

Example 7.5.76 *“King Zlatan would flip him if he ever gets one of his players working under pepito the coward who wants to break up Spain. So in short. Pepito is lying.”* (source:UK-RW-SPR)

Example 7.5.77 *“With gay people, the goal was to remove legal barriers or allow them to enjoy the legal benefits of marriage, such as inheritance, child custody. It was, in short, ‘We want what you have.’ Transgender issues by contrast, have moved beyond, ‘Don’t discriminate against us.’ to twisting language.”* (source:USA-RW-SPR)

In both the examples mentioned above, the commenters used ‘in short’ to summarize an idea. In the first comment from the UK sub-corpora, the commenter criticized Pepito (former footballer and coach), and in this regard, the commenter used it to mean something in a nutshell. In the second comment from the USA sub-corpora, the commenter showed the difference between the rights claimed by gay people and transgender people. Here, ‘in short’ was used by the commenter to summarize the rights of gay people. So, in both examples, the phrase ‘in short’ was used to say something briefly.

Announcers: The fourth lexico-grammatical category of *frame markers* is announcers. The writers use announcers to announce discourse goals. They were used by the writers to signal the direction of the argument and to provide an idea to the readers about the progression of the text by announcing future actions. In our study, the common occurrence we found in the three sub-corpora was ‘will’. It was enormously used

by the commenters of India, the UK, and the USA; however, the India commenters used it comparatively more (313.76 per 10,000 words), and the UK and USA commenters used it almost similarly (218.87 and 216.71 per 10,000 words, respectively). We found no other announcer in the India sub-corpora. The other announcers that we found from the UK and USA sub-corpora were *shown*, *seek*, and *aim to* which were less frequently used compared to ‘will’ by the commenters of the UK and the USA. Here, we mention some examples of ‘will’ from our corpus.

Example 7.5.78 *“I know this is going a bit off-topic, but let’s face it, the experience of watching a film at the cinema will always be more satisfying than watching it on TV via a streaming service.” (source:UK-LW-ENT:1)*

Example 7.5.79 *“Once you Pakistanis stop sending terrorists to India for an extended time then maybe we Indians will think about any bilateral sport activity. Until then keep barking and we Indian will focus on improving our economy, education, and living standards. Once Pakistani people wake up and realize this and give up supporting terrorists then let us know.” (source:IND-LW-SPR)*

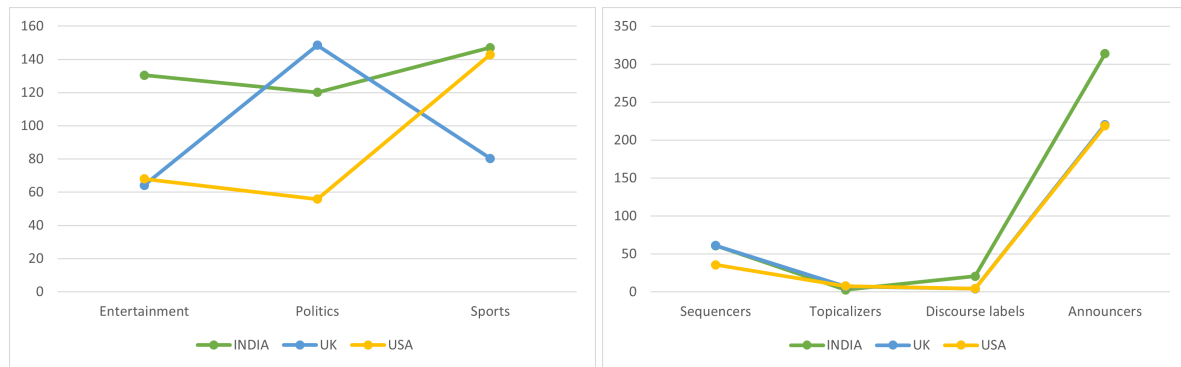
Example 7.5.80 *“Republicans made their pick: Trump. I’m sure the next Democratic president will compromise at least as much as Trump compromises with Democrats. It is not the duty of Democratic voters to embrace Republican vales.” (source:USA-LW-ENT)*

In the first example from the UK sub-corpora, the commenter used the announcer ‘will’ to state that watching a film at the cinema was more pleasing than watching it on TV. The commenter here expressed his view and tried to convince the other commenters about something that was going to occur in the future. In the second example from the India sub-corpora, the commenter used the announcer ‘will’ to refer to future action. The commenter revealed that to continue the bilateral sports activity between India and Pakistan, first, Pakistan needs to stop sending and sponsoring terrorists to

India. In this regard, the commenter used ‘will’ to announce something that has a probability to happen in the future. In the third example from the USA sub-corpora, the commenter sarcastically criticized Donald Trump (former President of the USA), and in this regard, he used ‘will’ to indicate a future action that is going to be taken by the next Democratic President.

Table 7.14: Top-5 (normalised frequency per 10,000 words) words of lexico-grammatical categories of *frame markers*

	INDIA		UK		USA	
Sequencers	first	51.74	first	27.56	first	18.08
	second	7.98	second	26.94	third	8.87
	finally	1.08	third	2.93	second	6.55
			next	2.80	to begin with	1.70
			finally	1.00	finally	0.45
Topicalizers	in terms of	2.65	regarding	3.57	regarding	6.96
	regarding	0.00	in terms of	1.93	in terms of	0.78
	with respect to	0.00	with respect to	1.17	with respect to	0.00
	with regard to	0.00	with regard to	0.47	with regard to	0.00
Discourse labels	in all	14.02	in short	3.12	briefly	2.06
	overall	4.67	overall	1.00	in short	1.52
	conclude	1.81			conclude	0.45
				in all	0.45	
Announcers	will	313.76	will	218.87	will	216.71
			show	1.67	seek	1.67
				aim to	0.45	



(a) Based on domains

(b) Based on lexico-grammatical categories

Figure 7.13: Comparisons of the occurrences (normalized) of *frame markers*

Comparisons of the occurrences of *frame-markers*

From figure 7.13a and 7.13b, we have drawn some noteworthy characteristics of using *frame markers* by the UK, India, and USA commenters.

- From figure 7.13a it can be seen that the pattern of using *frame markers* by the India and USA commenters was the same; however, in the case of UK commenters, it was entirely the opposite. Concerning the domains, it is evident that in the India and USA sub-corpora, frame markers were used in the following order: politics, entertainment, and sports. On the contrary, the UK commenters used them more in the politics domain, followed by the sports and entertainment domains. Interestingly, the use of *frame markers* was almost similar (on the lower side) in the entertainment domain of the UK and USA sub-corpora. This suggests that the UK and the USA commenters of the entertainment domain were not so eager to introduce different functions of the comments to the readers and commenters. One possibility could be that they want to leave the commenters to decide by themselves about the different functions within the comments.
- Among the lexico-grammatical categories, we found that the usage pattern was identical for the three countries. Sequencers were used by the India and UK commenters almost the same, whereas the USA commenters used them a bit less. Interestingly, topicalizers were used almost the same by the UK, India, and USA commenters. Another interesting fact was discourse labels and announcers were used almost the same in the UK and USA sub-corpora. Announcers were used most frequently by the India commenters which points out the tendency of the India commenters to announce future actions to the readers and other commenters and help them to understand the progression of the argument.

7.5.8 Interactive: Endophoric markers

The third category of *interactive* metadiscourse markers is *endophoric markers*. *Endophoric markers* are used by writers to refer to the other important parts of the text. *Endophoric markers* draw the reader's attention to additional information and make it available to them so they can reconstruct the writer's intentions by referring to other passages in the text Hyland (2010). Writers use *endophoric markers* to make explicit

references to other parts of the text (Hyland (2005a)). Previously, various authors, such as, Hyland (2005a), Cao and Hu (2014), Mur-Dueñas (2011) have studied endophoric markers as expressions that refer to the other parts of the text (for example, ‘as noted above’, ‘see table 1’). In our study, the commenters used *endophoric markers* in a threefold way: 1) to refer to some essential parts of the comment. 2) to include some additional information about previous as well as future ideas or events. 3) to direct the readers’ or commenters’ attention. This section depicts the occurrences of *endophoric markers* found in India, the UK, and the USA sub-corpora and their lexicogrammatical categories. Although the frequency of the *endophoric markers* was on the lower side in our corpus, it is crucial to mention it as it will help to understand the pattern of using *endophoric markers* by the commenters irrespective of country, domain, and political ideology. From our corpus, we found the following lexicogrammatical categories of *endophoric markers*: anaphoric (*before, effect, above, earlier, previously, paper, so far*) and cataphoric (*following, below*). Table 7.15 represents the normalized frequencies of *endophoric markers* found in India, the UK, and the USA.

India: The quantitative analysis of the lexicogrammatical categories of *endophoric markers* illustrates that the commenters of the politics domain used *endophoric markers* comparatively more, followed by the entertainment and sports domain. Surprisingly, the India commenters, irrespective of domain and political ideology, did not use cataphoric markers in their comments. Anaphoric markers were used significantly more in the left-wing sub-corpora (6.49 per 10,000 words) than in the right-wing (1.81 per 10,000 words). In the entertainment and sports domain, anaphoric markers were used only by the left-wing commenters for 2.77 and 2.23 per 10,000 words, respectively, which is almost the same. The right-wing commenters of these two domains did not use them a single time. The probable reason for not using cataphoric markers by the right-wing commenters could be that they were not interested in adding some additional information for the readers about future events. Possibly they wanted their readers and commenters to interpret the actions of an argument by themselves.

The UK: The UK commenters of the politics domain used *endophoric markers* comparatively more, followed by the sports and entertainment domain. In the politics domain, anaphoric markers were used almost 2 times more by left-wing commenters compared to right-wing ones. In the use of cataphoric markers, we saw that there was not a single use of them in the left-wing sub-corpora, whereas they were used significantly less in the right-wing sub-corpora (0.47 per 10,000 words). In the sports domain, the right-wing commenters did not use anaphoric and cataphoric markers; however, they were used for 3.99 and 1.00 per 10,000 words, respectively, by the left-wing commenters. In the entertainment domain, anaphoric markers were used only by the right-wing commenters, whereas they were not used a single time by the left-wing commenters which shows the left-wing commenters' disinterestedness in advising the readers and commenters about the past events. Both the left-wing and right-wing commenters did not use cataphoric markers in their comments. Perhaps, the commenters of both the left-wing and right-wing did not perceive the need to guide the readers and commenters by adding any additional information.

The USA: *Endophoric markers* were not found in the entertainment domain of the USA sub-corpora. Among the two remaining domains, the commenters of the sports domain overall used *endophoric markers* almost 3 times more in their comments than in the politics domain. Anaphoric markers were never used by the left-wing commenters of the sports domain, whereas the right-wing commenters used them for 4.31 per 10,000 words. The same probable reason could be thought of that the right-wing commenters of the sports domain let the readers and commenters decide the interpretation of the comments, without informing them about the previous events. In the use of cataphoric, the left-wing commenters used it a bit more than the right-wing. In the politics domain, cataphoric markers were never used by both the left-wing or right-wing commenters. Although the use of anaphoric markers was on the lower side, we found that they were used a bit more in the left-wing than in the right-wing sub-corpora.

Occurrences in Lexico-Grammatical categories: Here, we illustrate the in-

Table 7.15: Normalized frequencies of *endophoric markers* found in India, the UK, and the USA

		Entertainment			Politics			Sports		
		LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
INDIA	Anaphoric	2.77	0.00	2.77	6.49	1.81	8.30	2.23	0.00	2.23
	Cataphoric	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Total	2.77	0.00	2.77	6.49	1.81	8.30	2.23	0.00	2.23
UK	Anaphoric	0.00	3.85	3.85	4.66	2.34	7.00	3.99	0.00	3.99
	Cataphoric	0.00	0.00	0.00	0.00	0.47	0.47	1.00	0.00	1.00
	Total	0.00	3.85	3.85	4.66	2.81	7.47	4.98	0.00	4.98
USA	Anaphoric	0.00	0.00	0.00	1.14	0.89	2.03	0.00	4.31	4.31
	Cataphoric	0.00	0.00	0.00	0.00	0.00	0.00	1.69	1.08	2.77
	Total	0.00	0.00	0.00	1.14	0.89	2.03	1.69	5.39	7.08

depth analysis of the occurrences from each lexico-grammatical category of *endophoric markers*.

Anaphoric: The first lexico-grammatical category of *endophoric markers* is anaphoric, which is used by the writers (in our case, the commenters) to refer to some previous ideas or events that took place. The three anaphoric markers which were common among the three sub-corpora were ‘before’, ‘above’, ‘effect’, and ‘earlier’. ‘Before’ were used more frequently in the UK sub-corpora, followed by the USA and India sub-corpora. ‘Above’ was used more frequently by the UK commenters (10.28 per 10,000 words), followed by India (6.55 per 10,000 words) and the USA commenters (4.88 per 10,000 words). ‘Effect’ was used comparatively more by the India and UK commenters than the USA commenters. ‘Earlier’ was used comparatively more in the India sub-corpora, followed by the UK and USA sub-corpora. The other occurrences that we found from our corpus were ‘so far’, ‘paper’, ‘previously’. Here, we show some examples of ‘before’ from our corpus.

Example 7.5.81 “*GaGa may be good at publicising herself, but she doesn’t really have much to offer musically that hasn’t been done before.*” (*source:UK-LW-ENT:2*)

Example 7.5.82 “*Very jubilant to see to see our tri colour flying high. Congrats to whole contingent for such a never before performance.*” (*source:IND-RW-SPR*)

Example 7.5.83 “*Know your history before you speak. Bush gave tax incentives for*

American companies to shift manufacturing overseas. That was the beginning of our manufacturing decline.” (source:USA-LW-POL)

The respective commenters from the UK, India, and USA used ‘before’ in the three examples mentioned above to refer to some previous events that have already taken place. In the first example from the UK sub-corpora, the commenter used ‘before’ to indicate something that had taken place earlier. The commenter criticized Lady Gaga (an American singer), and in this regard, the commenter stated that the contribution of Lady Gaga was not as significant as that has not been done by the other singers earlier. In the second example from the India sub-corpora, the commenter used ‘before’ to refer to the performance that had not been done earlier by the Indian players in the Commonwealth Games. In the third example from the USA sub-corpora, the commenter advised his fellow commenter to have the necessary knowledge about the past before commenting on some issue, and in this regard, the commenter used ‘before’ to refer to the previous events or the history.

Cataphoric: The second lexico-grammatical category of *endophoric markers* is cataphoric, which the writers use to refer to the following idea or to mention some event that will take place in the future. Interestingly, we have not found any cataphoric in the India sub-corpora; however, we found cataphoric significantly less in the UK and USA sub-corpora. We found two cataphoric markers from the UK and USA sub-corpora: ‘following’, and ‘below’. They were found to be used by the commenters of the UK and USA; however, they were used significantly less in both the sub-corpora. The UK and USA commenters used ‘following’ almost similarly, whereas another cataphoric ‘below’ was used comparatively more in the USA sub-corpora than in the UK. Here, we present some examples of ‘below’ from our corpus.

Example 7.5.84 *“Richard Simian writes, below : I am not sure if these people will EVER stop bleating, we need to get on with making it work. I am thoroughly ashamed of half of this country and their ‘can’t do’ attitude. We CAN make Brexit a success and we WILL do. Get behind YOUR country and stop this EU cult worship. Well... I’m*

thoroughly proud of that half of the country which has the intelligence and the strength of character NOT to pile in unthinkingly behind a thoroughly bad idea...” (source:UK-RW-POL:2)

Example 7.5.85 “Go get em King James...alabama is really not use to strong free black people ‘talking back’ to their football hero/coach so expect the normal covert bigotry to emerge like in a few of these comments below...” (source:USA-LW-SPR)

In the first example from the UK sub-corpora, the commenter referred to the utterance of a person that was mentioned beneath. He used the preposition ‘below’ to indicate the direction where the utterance of the commenter was written. In the second example from the USA sub-corpora, while conversing with one of his fellow commenters, the commenter used the preposition ‘below’ to refer to the direction where, according to him, some expected biased comments will come from some of the commenters. Similar to the first comment, it was used to refer to something that occurred beneath in the same comment.

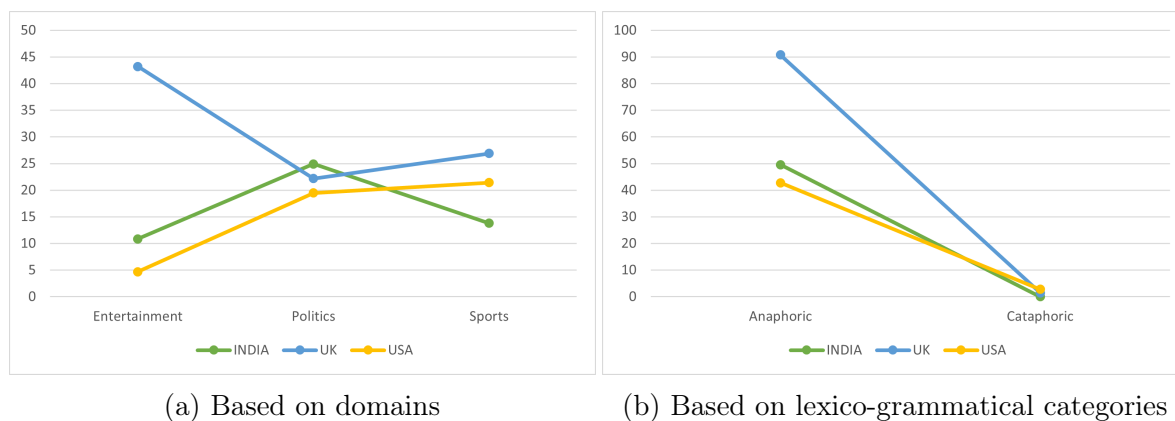


Figure 7.14: Comparisons of the occurrences (normalized) of *endophoric-markers*

Comparisons of the occurrences of *endophoric markers*

Figure 7.14a and 7.14b depicts some notable characteristics of using *endophoric markers* by the UK, India, and USA commenters.

Table 7.16: Top-5 (normalized frequency per 10,000 words) words of lexico-grammatical categories of *endophoric markers*

	INDIA		UK		USA	
Anaphoric	before	24.34	before	65.02	before	30.64
	effect	11.94	effect	10.95	effect	5.79
	above	6.55	above	10.28	above	4.88
	earlier	3.85	earlier	0.94	paper	0.57
	so far	1.81	previously	0.47	earlier	0.45
Cataphoric			following	1.00	following	1.08
			below	0.47	below	1.69

- From figure 7.14a, it is evident that the usage pattern concerning *endophoric markers* was different for the three domains of the UK, India, and the USA. The India commenters used *endophoric markers* most in the politics domain. In contrast, the UK commenters used them more frequently in the entertainment domain, and the USA commenters used them more in the sports domain. So, the different patterns of using *endophoric markers* by the commenters of the UK, India, and the USA suggests that they think differently while guiding the readers and commenters about the essential parts of the comments or adding some additional information about the past or future events. *Endophoric markers* were used less frequently in the entertainment domain by the India and USA commenters; however, the UK commenters used them less frequently in the politics domain. The UK commenters overall used them more than the India and USA commenters.
- Figure 7.14b depicts the usage pattern of the lexico-grammatical categories of *endophoric markers* in the UK, India, and USA sub-corpora. It can be clearly noticed that the usage pattern was the same, although there was a difference in the frequency of use by the commenters of the respective nations. Anaphoric markers were used more or less the same in the India and USA sub-corpora; however, in the UK sub-corpora, they were used comparatively more. In using cataphoric markers, it was seen that the UK and USA commenters used them significantly less, while the India commenters never used them once in their com-

ments One possibility of this could be that the India commenters were not so willing to direct the readers and commenters about specific actions within the comments so that the readers could get a chance to interpret the actions by themselves.

7.5.9 Interactive: Code Glosses

The fourth category of *interactive* metadiscourse markers is *code glosses* which the writers use to explain clearly what has been stated by the writer before. These markers help the readers to understand the appropriate meaning of the text. They also help the writers to rephrase, elaborate and explain ideas or subjects to make the readers understand their meaning. The writers used *code glosses* to guide the readers by clarifying some crucial points where the readers need help to understand or to add some examples for a better understanding of the idea or the text. Previously, various authors, such as, Hyland (2005a), Hyland (2007), Dehghan and Chalak (2016) have studied *code glosses*. In our study, *code glosses* were used by the commenters in three ways: 1) to include further information, 2) to reformulate or rephrase, 3) to explain different actions, and 4) to provide examples of what comes before. *Code glosses* include two lexico-grammatical categories: a) reformulation and b) exemplification. From our corpus, we got the following lexico-grammatical categories and their occurrences of *code glosses*: Reformulation (*finally, that is, rather, : (paranthesis), in other words, this means, i.e., in particular, specifically, namely*) and Exemplification (*as in, such as, for example, example, : (colons), for instance, e.g.*). In this section, we explain the lexico-grammatical categories of *code glosses* and their occurrences from the UK, India, and the USA sub-corpora. Table 7.17 depicts the normalized frequencies of *code glosses* found in the India, UK, and USA sub-corpora.

India: The quantitative analysis of *code glosses* from the India sub-corpus shows the fact that the commenters of politics domain used *code glosses* more frequently, followed by the sports and entertainment domains. In the politics domain, both reformulation and exemplification were used comparatively more by the right-wing commenters compared to the left-wing ones which shows the tendency of the right-wing commenters of the politics domain of explaining the actions of the comments to the readers and commenters by including more information. In the sports domain, the overall use of *code glosses* was seen more in the left-wing sub-corpora. However, reformulation

was used comparatively less in the left-wing sub-corpora than in the right-wing. The right-wing commenters of the sports domain did not use exemplification, whereas the left-wing commenters used it for 6.68 per 10,000 words. In the entertainment domain, reformulation was used more frequently by left-wing commenters; however, right-wing commenters never used reformulation for once. In the use of exemplification, we found that the left-wing and the right-wing commenters used them almost the same, and they were on the lower side.

The UK: Interestingly, if we look at the overall use of *code glosses* by the UK commenters, we can see that they were used almost the same in the politics and sports domain and comparatively less in the entertainment domain. In the politics domain, reformulation was used more in the right-wing sub-corpora than the left-wing. On the contrary, exemplification was used almost more than 2 times in the left-wing sub-corpora than in the right-wing sub-corpora. In the sports domain, the right-wing commenters used both reformulation and exemplification comparatively more than the left-wing. In the entertainment domain, reformulation was used almost more than 2 times in the right-wing sub-corpora. In the use of exemplification, we observed that the use was on the lower side by the left-wing, whereas the right-wing never used them a single time. Although the left-wing commenters of the entertainment domain used exemplification very less, it shows the unwillingness of both the left-wing and right-wing commenters to not explain to the readers and commenters the actions within the comments.

The USA: The categories of *code glosses* were used most frequently in the sports domain, followed by the politics and entertainment domain. In the sports domain, reformulation was used comparatively more in the right-wing sub-corpora. Exemplification was used by the left-wing commenters significantly less (3.39 per 10,000 words), while the right-wing commenters used them almost 6 times more (19.40 per 10,000 words). In the politics domain, reformulation was significantly used more in the right-wing sub-corpora than the left-wing one. Exemplification was used significantly less in

the left-wing sub-corpora (1.70 per 10,000 words), while they were used almost 8 times more in the right-wing sub-corpora (9.37 per 10,000 words). In the entertainment domain, reformulation was used comparatively more in the right-wing sub-corpora; however, exemplification was used only by the left-wing, whereas they were not used once by the right-wing. The reason for using reformulation by both the left-wing and right-wing commenters indicates the willingness of these commenters to include further information in the comments by rephrasing or reformulating the actions within the comments.

Table 7.17: Normalized frequencies of *code-glosses* found in India, the UK, and the USA

		Entertainment			Politics			Sports		
		LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
INDIA	Reformulation	8.31	0.00	8.31	8.65	9.06	17.71	4.45	9.35	13.80
	Exemplification	2.77	2.65	5.43	4.33	7.25	11.57	6.68	0.00	6.68
	Total	11.09	2.65	13.74	12.98	16.31	29.28	11.13	9.35	20.47
UK	Reformulation	10.00	23.09	33.09	15.16	20.12	35.28	19.93	21.23	41.15
	Exemplification	3.33	0.00	3.33	12.82	5.15	17.97	3.99	7.96	11.94
	Total	13.33	23.09	36.43	27.98	25.27	53.25	23.91	29.19	53.10
USA	Reformulation	15.58	26.74	42.32	16.47	33.00	49.47	22.02	30.17	52.20
	Exemplification	9.35	0.00	9.35	1.70	9.37	11.07	3.39	19.40	22.78
	Total	24.92	26.74	51.66	18.17	42.37	60.54	25.41	49.57	74.98

Occurrences in Lexico-Grammatical categories: Here, we describe each occurrence from the lexico-grammatical categories of *code glosses* from the India, UK, and USA sub-corpora.

Reformulation: The first lexico-grammatical category of *code glosses* is reformulation, which is used by writers to restate a particular idea in a different way that has been stated before for better understanding. The most commonly used reformulation from our corpus was ‘:’ (parenthesis), ‘that is’, and ‘finally’. We noticed an extensive amount of parenthesis (:) was used in the three sub-corpora where the USA commenters used it most frequently, followed by the UK and India commenters. The phrase ‘that is’ was used extensively by the UK and USA commenters (39.25 and 38.03 per 10,000 words, respectively), whereas the India commenters used it significantly less (10.87 per 10,000 words). The adverb ‘finally’ was used comparatively more in the USA

sub-corpora followed by the India and UK sub-corpora. There were some occurrences that were common only in the UK and USA sub-corpora: ‘rather’, ‘in other words’, and ‘i,e’. Here, we cite some examples from the UK, India, and the USA sub-corpora.

Example 7.5.86 *“We will have three options : Exit from Brexit Continue to share Sovereignty with the EU Suffer some financial loss from Brexodus that has already occurred. A Cameron++ deal is always on the table and the Champagne on ice. Accept the E27 offer...” (source:UK-LW-POL:2)*

Example 7.5.87 *“nothing adds up : They were at the wedding, ok good. Sri Devi stays back in UAE while husband and daughter returns to Mumbai. Why did she stay back w/o her husband. Then husband returns to ‘surprise’ her at 5.30pm. Chatted for 15 mins. Then bath time and then dinner outing. She was asleep during day time and alone. Would she be still alive.” (source:IND-LW-ENT)*

Example 7.5.88 *“If you have a team that gives up and acknowledged that there is a team they can’t compete with, then it’s time to break it down and start over. Start with a new GM and build a team with the athletes and shooters that can compete. If Wiz don’t do that, here are teams they will not be able to compete with next year: Hornets, Pacers, Heat, Celtics, Magic, Raptors and Cavaliers. Thats just here in the East. I hope the Owner knows enough about basketball to realize that.” (source:USA-LW-SPR)*

In the examples mentioned above from the three sub-corpora, parenthesis (:) were used by the respective commenters to either include further information or to reformulate. In the first comment from the UK sub-corpora, the commenter included three options that he thought to be suitable after Brexit, among which the UK could select one. Here, he used the parenthesis (:) to include further information. In the second comment by the Indian commenter, parenthesis (:) was used to add information about the incidents that took place at the time of Sridevi’s (a famous Indian film actress) death. The commenter here listed the incidents, and in this regard, parenthesis (:) was used.

was used by the commenter to list the events. In the third example from the USA sub-corpora, the commenter used parenthesis (:) to add the names of the basketball teams. However, it was interesting that both the commenters from the UK and India sub-corpora never used enumeration to list the options and events; instead, the commenters included short sentences one after another to list the events.

Exemplification: The second lexico-grammatical category of *code glosses* is exemplification which is used by the writers to provide examples of what comes before. They are used by the writers to make the idea more concrete by providing examples. The four common exemplifications that we found from the three sub-corpora were: ‘example’, ‘such as’, ‘for example’, and ‘as in’. The USA commenters used ‘example’, ‘such as’, and ‘as in’ most frequently in their comments, followed by the UK and India commenters. The commenters of India, the UK, and the USA used ‘for example’ almost the same. The other occurrences which were common in the UK and USA sub-corpora were ‘:’ (colons), ‘for example’, and ‘for instance’. Here, we bring some examples of the phrase ‘such as’ from the UK, India, and the USA sub-corpora.

Example 7.5.89 *“The Lisbon Treaty was changed in an attempt to assuage opposition. For example, it removed references to EU symbols such as the flag, anthem, motto, currency and ‘Europe Day’ that had given rise to fears that a ‘superstate’ was being created.” (source:UK-RW-POL:1)*

Example 7.5.90 *“Shooters should not bother. There are other games such as Asian, the Olympic etc where they can contest and where competition is better than at CW games. Win the medals at the Olympics and then we will cheer more.” (source:IND-LW-SPR)*

Example 7.5.91 *“I’m looking forward to male-born, high school seniors who suddenly discover that they’re women and demanding admission to all women’s colleges and bastions of feminism such as Smith, Barnard, and Mt. Holyoke. Of course, the-newly transgendered women don’t have to wear long hair, makeup, or dresses, which are social constructions anyway.” (source:USA-RW-SPR)*

In the above-mentioned examples from the UK, India, and the USA sub-corpora, the respective commenters used the phrase ‘such as’ to include some examples. In the first comment from the UK sub-corpora, the commenter used ‘such as’ to include some symbols that stand for the EU (European Union). In the second comment from the India sub-corpora, the commenter used ‘such as’ to include various international events where according to the commenter, the Indian shooters will get a chance to perform and win as well. In the third comment from the USA sub-corpora, the commenter used the phrase ‘such as’ to include the names of some famous feminists (the believers in social, economic, and political equality for women), as examples.

Table 7.18: Top-5 (normalised frequency per 10,000 words) words of lexico-grammatical categories of *code glosses*

	INDIA		UK		USA	
Reformulation	: (parenthesis)	21.29	: (parenthesis)	54.06	: (parenthesis)	78.99
	that is	10.87	that is	39.25	that is	38.03
	finally	7.66	finally	5.29	finally	9.21
			rather	5.26	specifically	3.94
			in other words	2.10	i.e.	3.36
			i.e.	1.93	in other words	2.81
			this means	1.63	rather	2.24
					in particular	2.09
					this means	1.86
					namely	0.45
Exemplification	example	11.40	example	13.96	example	16.06
	such as	7.28	such as	8.06	such as	12.16
	for example	2.77	as in	4.47	: (colons)	5.35
	as in	2.23	: (colons)	3.50	as in	4.91
			for example	2.10	for example	2.75
			for instance	1.17	for instance	1.52
					e.g.	0.45

Comparisons of the occurrences of *code glosses*

From figure 7.15a and 7.15b, we have drawn some significant characteristics of using *code glosses* by the commenters of the UK, India, and the USA.

- Figure 7.15a portrays the domain wise usage pattern of *code glosses* that we found from the three sub-corpora. Among the three domains, irrespective of country, the commenters used *code glosses* less frequently in the entertainment domain. In the UK and India sub-corpora, the commenters used them more frequently in

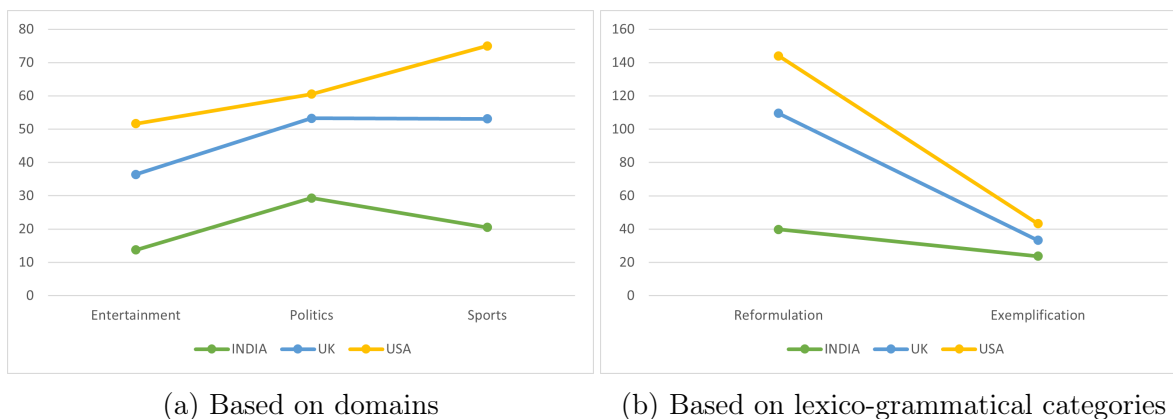


Figure 7.15: Comparisons of the occurrences (normalized) of *code-glosses*

the politics domain; however, in the USA sub-corpora, they were used most in the sports domain. Interestingly, in the UK sub-corpora, the use of *code glosses* was almost the same in the politics and sports domains which hints at the fact that the commenters of both the politics and sports domain tended to add information and tried to present the actions within the comments more clear to the readers and commenters.

- From figure 7.15b it is evident that the pattern of using the lexico-grammatical categories of *code glosses* by the commenters of UK, India, and USA sub-corpora was the same. Irrespective of political ideology and domains, the commenters from the UK, India, and the USA used reformulation comparatively more in their comments compared to exemplification. It was observed that in using exemplification, the India commenters used them less frequently, while the USA commenters used them most frequently (among the three sub-corpora) in their comments. One possible explanation could be advanced here that the USA commenters were more into providing examples to give a clear picture to the readers and commenters.

7.5.10 Interactive: Evidentials

The final category of *interactive* metadiscourse markers is *evidentials*. *Evidentials* are used by the writers to refer to the source of information from other texts (Hyland (2005a)). According to Hyland (2010, p. 129), *evidentials* “indicate the source of textual information which originates outside the current text.” It involves two lexicogrammatical categories: personal and impersonal. As we have mentioned before, in our study, we have followed the metadiscourse model proposed by Hyland (2005a), but still, we found some differences in our corpus regarding *evidentials*. Mostly in written academic discourse, citations are perceived as *evidentials*; however, as our study is on digital comments, direct citations were not used by the commenters. Instead, we found that the commenters used *evidentials* while mentioning a particular reference within the comment itself. In our corpus, *evidentials* were the least used interactive markers by the commenters. From the three sub-corpora, we found that *evidentials* were used by the commenters significantly less. We found the following lexicogrammatical categories of *evidentials* and their occurrences: Personal (*following, consequence, approach*) and Impersonal (*model, research, earlier, previous*).

India: The overall use of *evidentials* was more in the politics domain, followed by the sports and entertainment domains. In the politics domain, personal markers were used in the left-wing sub-corpora for 3.24 per 10,000 words; however, they were never used a single time in the right-wing sub-corpora. On the contrary, with the use of impersonal markers, only the right-wing commenters used them for a little bit (1.81 per 10,000 words). The left-wing commenters did not use them once. Personal markers were not used in the sports domain, irrespective of political ideology. A possible reason for this could be that the India commenters of the sports domain did not perceive the need to mention the source of information before stating them. Impersonal markers were used in the left-wing sub-corpora, while they were never used in the right-wing sub-corpora. Personal markers were used for 2.77 per 10,000 words in the entertainment domain, whereas they were never used in the right-wing sub-corpora. Irrespective

of political ideology, impersonal markers were never used by the commenters of the entertainment domain.

The UK: The quantitative analysis of *evidentials* in the UK sub-corpora portrays the fact that *evidentials* were used more frequently in the politics domain, followed by the sports and entertainment domains. In the politics domain, personal markers were used almost the same in both the left-wing and right-wing sub-corpora which shows the tendency of the commenters of the sports domain to refer to the source of information to validate their stance within the comments. The right-wing commenters of the politics domain used impersonal markers for 3.74 per 10,000 words, whereas they were never used in the left-wing sub-corpora. In the sports domain, the left-wing commenters used both personal and impersonal markers very less; however, both these markers were never used in the right-wing sub-corpora which depicts the unwillingness from the commenters' side to authenticate the statements within the comments. In the entertainment domain, irrespective of political ideology, personal markers were never used. In the right-wing sub-corpora of the entertainment domain, the UK commenters used impersonal markers for 3.85 per 10,000 words, while they were never used in the left-wing sub-corpora.

The USA: In the use of *evidentials* in the USA sub-corpora, we found that they were used more in the sports domain, followed by the politics and entertainment domains. In the sports domain, personal and impersonal markers were used comparatively less in the left-wing sub-corpora, whereas in the right-wing sub-corpora, they were used almost 4 times and 3 times more, respectively. In the politics domain, personal markers were used less in the left-wing sub-corpora, while they were used comparatively more in the right-wing. On the contrary, the commenters of the left-wing sub-corpora used impersonal markers comparatively more than the right-wing. In the entertainment domain, the left-wing commenters used *evidentials* almost 2 times more than the right-wing. However, we found that irrespective of political ideology, no impersonal markers were used by the commenters of the entertainment domain of the USA which

shows the reluctance of the commenters to refer to the shared information within the comments.

Table 7.19: Normalized frequencies of *evidentials* found in India, the UK, and the USA

		Entertainment			Politics			Sports		
		LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
INDIA	Personal	2.77	0.00	2.77	3.24	0.00	3.24	0.00	0.00	0.00
	Impersonal	0.00	0.00	0.00	0.00	1.81	1.81	4.45	0.00	4.45
	Total	2.77	0.00	2.77	3.24	1.81	5.06	4.45	0.00	4.45
UK	Personal	0.00	0.00	0.00	2.33	2.34	4.67	2.99	0.00	2.99
	Impersonal	0.00	3.85	3.85	0.00	3.74	3.74	1.99	0.00	1.99
	Total	0.00	3.85	3.85	2.33	6.08	8.41	4.98	0.00	4.98
USA	Personal	4.67	2.06	6.73	1.70	2.23	3.93	1.69	4.31	6.00
	Impersonal	0.00	0.00	0.00	2.84	0.89	3.73	1.69	3.23	4.93
	Total	4.67	2.06	6.73	4.54	3.12	7.67	3.39	7.54	10.93

Occurrences in Lexico-Grammatical categories

In this section, we will discuss each occurrence from the lexico-grammatical categories of *evidentials*. These occurrences were from the UK, India, and USA sub-corpora. Table 7.20 outlines the occurrences of the lexico-grammatical categories of *evidentials* from our corpus.

Personal: Personal markers are used by the writers to refer to the source information from other texts. In our study, we noticed that there was no such reference to the source information by the commenters; instead, the commenters used a particular reference within the comments by mentioning directly its presence only. Two personal markers were common in the three sub-corpora: ‘following’ and ‘consequence’. Another personal marker ‘approach’ was used in the UK and USA sub-corpora only. ‘Following’ were used comparatively more in the India sub-corpora, while they were used almost the same in the UK and USA sub-corpora. ‘Consequence’ was used comparatively more in the USA sub-corpora compared to the UK and India sub-corpora. In the use of ‘approach’, we saw that the UK commenters used them less frequently than the USA commenters. The following examples give a clear view of personal markers from our corpus.

Example 7.5.92 *“It has started already. The Chartered Institute of Procurement and*

Supply has found that one in seven EU companies with UK suppliers have moved part or all of their business out of Britain ‘to reduce their exposure to any complications resulting from Brexit’. The institute warned of “an imminent collapse in the UK’s supply chain following Brexit.” (source:UK-RW-POL:2)

Example 7.5.93 *“ITS PITY THAT SYMBOLS OF SLAVERY, which include, following islamic cult, wearing burqa/hizab, skull cap, reading quran, namaz, saying m, growing mustache-less beard, giving islamic name to children, and praise/defend those islamic terrorists etc etc, are STILL continued to be followed by our Hindu-converts!...All these are Reminders of slavery, which converts were forced to go through during islamic terrorist rule in India-Pakistan-Bangladesh ... (source:IND-LW-POL)*

Example 7.5.94 *“Jimmy Carter was a great president and one of the best in the history of this country. Unlike your warmonger opinion, which obviously following Trump and Netanyahu, he was and still is a peace loving true human being.” (source:USA-LW-POL)*

In all three examples mentioned above, the respective commenters used ‘following’ to mean ‘using the same method or theory’. In the first example, the UK commenter used ‘following’ to refer to a situation coming after Brexit. In the second example from the India sub-corpora, the commenter described that the religiously converted people of India still now bear the burden of slavery by following various practices. Here, the commenter used ‘following’ to mean ‘using the same method’. In the third example from the USA sub-corpora, the commenter praised Jimmy Carter (former US President) and criticized his fellow commenter who was a supporter of Trump (the then US President). Here, the commenter used ‘following’ to mean observing or obeying a particular person. The above examples depict that the commenters referred to something in an implicit manner by mentioning the proposition within the comments.

Impersonal: Impersonal markers were used by the writers to refer to something already existing or something that appears to the writer to be common. Like the

personal markers, impersonal markers were also used very less by the commenters. From our corpus, we found one common impersonal marker ‘model’ in the three sub-corpora. They were used 3 times more in the India sub-corpora, while they were used less frequently in the UK and USA sub-corpora. The other occurrences from the UK and USA sub-corpora were ‘research’, ‘earlier’, and ‘previous’. Here, we mention some examples from the UK, India, and the USA sub-corpora that will help to understand the use of impersonal markers in our corpus.

Example 7.5.95 *“Should we have decided to follow the Norway model, I don’t know. Did most leave voters vote for this, as you assert. I don’t know and I wonder how you do. Have you asked each and every one. No the Norway model would suit you. To be frank it would suit me.”* (source:UK-RW-POL:2)

Example 7.5.96 *“We have stopped producing athletes altogether. What is wrong-is it no infrastructure or no motivation? This non-proactive, non-reactive Captain Government can, may be, study Haryana’s model of sport success and do some something. It almost feels like Sukhbeer-III. And then he will be back for the 4th Innings.”* (source:IND-LW-SPR)

Example 7.5.97 *“Just curious, Why do you blame China and not Walmart who created the business model, why not Walmart shoppers who would rather buy cheap for themselves than spend more for any ideal?”* (source:USA-LW-POL)

In the examples mentioned above from the three sub-corpora, the respective commenters used the noun ‘model’ to refer to a representation of a thing or a structure. The commenters also meant to refer to something which is pre-existing. In the first example, the commenter referred to a model named the Norway model (a model that includes two major European organizations), a proposal for a post-Brexit settlement. In this comment, the noun ‘model’ was used to refer to a particular structure or approach that should be followed. Similarly, in the second example, the Indian commenter advised

the-then Government to focus on Haryana’s model (a model that inspires the players to be successful and bring medals from International games, and instead of that, the Government will secure the future of those aspiring players). Similar to the first and second examples, in the third example from the USA sub-corpora, the commenter criticized Walmart’s business model (an American multinational retail corporation). In the second and third examples, the respective commenters used the noun ‘model’ to refer to an approach or a structure that should be followed. Therefore, from these examples it is evident that in our corpus *evidentials* were used a bit differently where the commenters mentioned the impersonal evidential markers directly in their comments instead of referring to information from other sources or texts.

Table 7.20: Top-5 (normalised frequency per 10,000 words) words of lexico-grammatical categories of *evidentials*

	INDIA		UK		USA	
Personal	following	3.85	consequence	5.26	consequence	11.59
	consequence	1.08	following	1.40	approach	3.28
			approach	1.00	following	1.79
Impersonal	model	6.26	research	4.78	research	5.50
			model	2.81	model	2.26
			earlier	1.99	previous	0.89

Comparisons of the occurrences of *evidentials*

From Figure 7.16, we extracted some notable characteristics of using *evidentials* in the comments by the commenters of the India, UK and USA.

- From Figure 7.16a, it is evident that regarding the domains the pattern of using *evidentials* was different. We can see that the UK commenters employed *evidential* markers in politics domain comparatively more than the commenters of USA and India. Irrespective of countries, *evidentials* were used less frequently in the entertainment domain. The USA commenters used *evidentials* most frequently in the sports domain. The India and UK commenters used *evidentials* almost the same in the sports domain. Overall, the India commenters used *evidentials* less

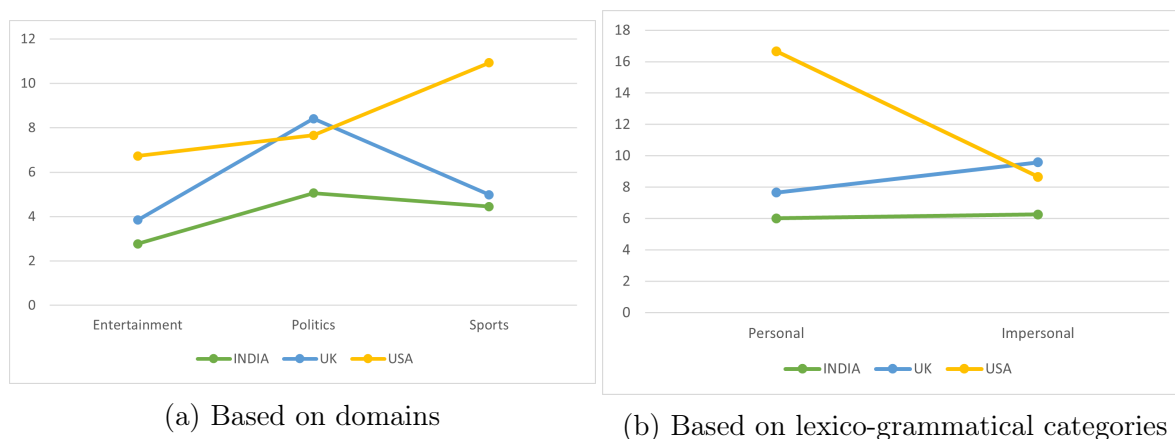


Figure 7.16: Comparisons of the occurrences (normalized) of *evidentials*

frequently, whereas the USA commenters used them most. Overall, the less frequent use of *evidentials* by the India commenters suggests the fact that as digital comments are informal in nature compared to the written academic discourse, there could be a possibility that the commenters did not find it necessary to refer to the source of information.

- Analysing our corpus based on the lexico-grammatical categories of *evidentials*, we conclude from Figure 7.16b that the pattern of using *evidentials* was almost the same for India and UK; however, it was opposite in the case of USA. In the India and UK sub-corpora, the commenters used personal markers less frequently and impersonal markers a bit more in their comments. It depicts the tendency of the India and UK commenters to share pre-existing information with the readers and commenters. In the USA sub-corpora, personal markers were used comparatively more, while the impersonal markers were used significantly less by the commenters. Interestingly, in the India sub-corpora, both personal and impersonal markers were used almost the same. Now, some possible explanations for the commenters' reluctance to use citations or references to the source information could be the followings: 1) News comments are usually short compared to written academic discourse. So maybe the commenters were not willing to write more. 2) The commenters preferred to just mention the source within

the conversation by only mentioning the names or topics. 3) As news comments are informal, the commenters were free to write anything without giving proper justification or reference, unlike academic texts.

7.6 Domain wise Analysis of Comments

The domain-specific analysis of using metadiscourse markers for each country is presented in Sections 7.2, 7.3, 7.4. In this section, we capture the quantitative analysis of metadiscourse markers with respect to domains in accordance with different political ideologies and nations as follows:

- Irrespective of the country, all the comments of a specific domain were put together to capture the quantitative analysis of domains based on the political ideologies of the commenters, in particular, left-wing and right-wing. In this regard, we present the use of interactive and interactional markers in Table A.7 and Table A.8 along with the distribution in Figure 7.17 and Figure 7.18.
- All the comments of a specific domain are analyzed based on the countries. According to the scope of this dissertation, we have analyzed three domains (namely- sports, politics, and entertainment) based on the three countries (namely- India, the UK, and the USA). The frequencies are reported in Table A.9, and Table A.10, and the distribution of the frequencies are depicted in Figure 7.19 and Figure 7.20.

The domainwise quantitative analysis of using *interactive metadiscourse markers* with respect to political ideology reveals that *transition markers* was the prime marker used across all the domains. In the sports domain, *transition markers* were used enormously, followed by the politics and entertainment domains.

The following *interactive marker* used most frequently was *frame markers*, which were used more in the sports domain compared to the politics and entertainment

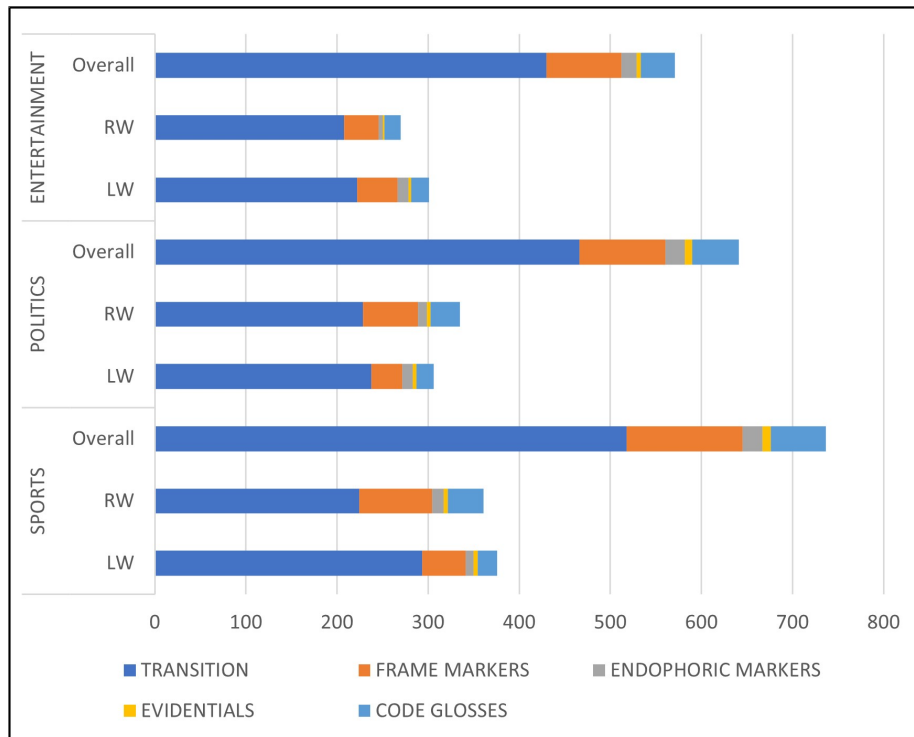


Figure 7.17: domain wise: distribution of *interactive markers* (normalized)

domains. The subsequent *interactive markers* used by the commenters with respect to domain were *code glosses*, *endophoric markers*, and *evidentials*. We found that irrespective of political ideology, *code glosses* were used comparatively more by the commenters than the *endophoric markers* which hints at the commenters' tendency to represent an idea or proposition in front of the readers and commenters by rephrasing and reformulating to make it more clear. The least used *interactive markers* by the commenters were *evidentials* which were used least in the entertainment domain (4.90 per 10,000 words), then in the politics domain (7.93 per 10,000 words) and sports domains (9.01 per 10,000 words), respectively.

All the *interactive markers* were used across all the domains. Among all the domains, *interactive markers* were used extensively in the sports domain in comparison to the politics and entertainment domains. In the sports domain, only *transition markers* were used more in the left-wing sub-corpora. In contrast, all the other *interactive markers* were used comparatively more in the right-wing sub-corpora. This demon-

strates that the right-wing commenters of the sports domain have a tendency to guide the readers and commenters through the propositions and arguments. In the politics domain, *transition markers* and *endophoric markers* were used more in the left-wing sub-corpora, while in the right-wing sub-corpora, *frame markers*, *evidentials* and *code glosses* were used comparatively more. Interestingly, in the entertainment domain, all the *interactive markers* were used comparatively more in the left-wing sub-corpora than the right-wing (See figure 7.17).

The domain wise quantitative analysis of using *interactional metadiscourse markers* with respect to political ideology represents the fact that *hedges* was the prime marker used across all the domains. In the sports domain, *hedges* were used extensively (610.32 per 10,000 words), followed by the entertainment (525.13 per 10,000 words) and politics domains (512.69 per 10,000 words). The succeeding *interactional marker* that was used most was *attitude markers* which were used most frequently in the sports domain, followed by the entertainment and politics domain. The next *interactive markers* mostly used were *engagement markers*, *boosters*, and *self-mentions*.

All the *interactional markers* were used across all the domains. Among all the *interactional markers*, *hedges* were used mostly in the sports domain, followed by the entertainment and politics domains. With respect to the domains, it was noticed that in the entertainment and politics domains *interactional markers* were used more by the right-wing commenters compared to the left-wing, except for *engagement markers* in the entertainment domain, which were used more by the commenters of the left-wing. However, we also noticed that in the sports domain, *hedges* and *self-mentions* were used comparatively more by the right-wing, whereas the other *interactional markers* were used more by the left-wing.

While comparing the *interactive* and *interactional* metadiscourse markers with respect to domains, we found that the *interactional markers* were used more (almost double) by the commenters in comparison to the *interactive markers* which shows the proneness of the commenters to include the readers and commenters in the arguments.

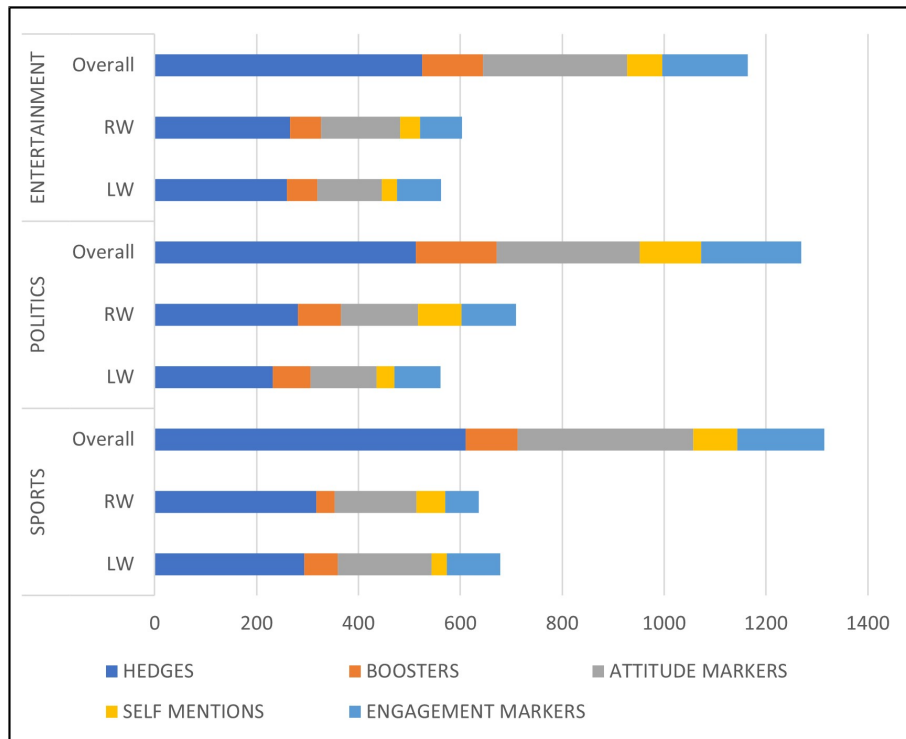


Figure 7.18: domain wise: distribution of *interactional markers* (normalized)

(See figure 7.17 and 7.18).

In the use of *interactive markers* based on domains, we found that the commenters from all three nations used *interactive markers* in their comments. For example, as it is evident from Figure 7.19, in the entertainment and politics domain, the UK commenters used *interactive markers* more compared to the other two nations: India and the USA, respectively. In the sports domain, the India commenters used *interactive markers* less in their comments compared to the USA and UK commenters. The UK commenters used *interactive markers* more in the sports domain. Among all the *interactive markers*, *transition markers* were used most frequently by the India, UK, and USA commenters. The succeeding *interactive marker* used by the commenters of three countries was *frame markers*. *Code glosses* and *endophoric markers* were used next by the commenters of India, the UK, and the USA. We noticed that *evidentials* were the least used *interactive markers* by the India, UK, and USA commenters. This suggests the disinterestedness of the commenters of all three nations to mention the source of the information in the

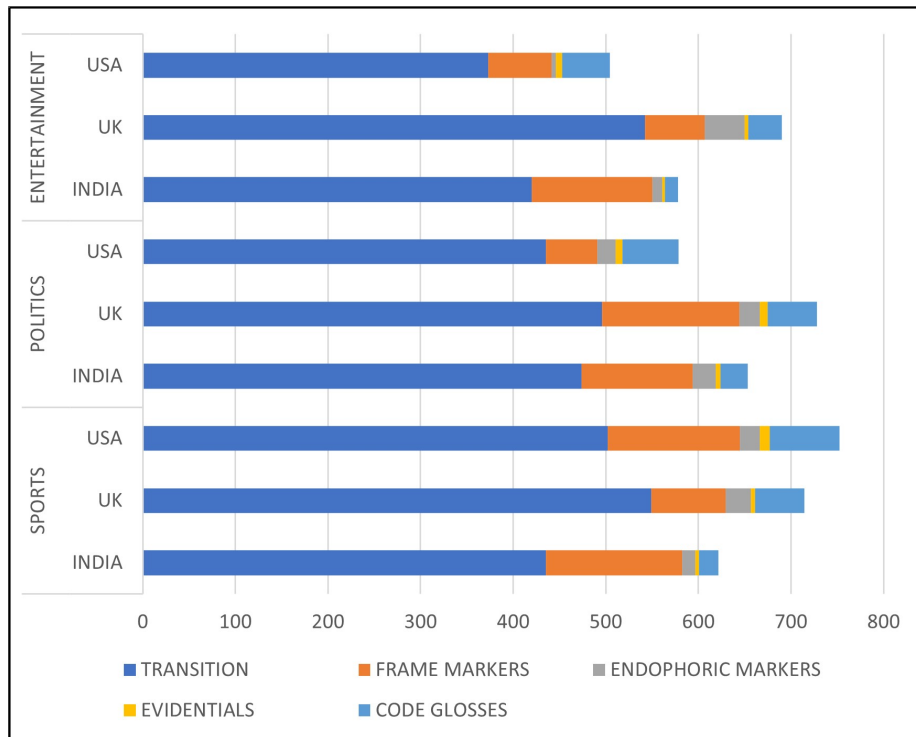


Figure 7.19: domain wise distribution (normalized) of *interactive markers* with respect to country

arguments.

In the use of *interactional markers* based on domains, we observed that in the entertainment domain, the India commenters used *interactional markers* most frequently, followed by the UK and the USA commenters. In the politics domain, the UK commenters used *interactional markers* most compared to the USA and India commenters. The USA and India commenters used *interactional markers* almost equally, with a slight difference in the politics domain. In the sports domain, the UK commenters used *interactional markers* most frequently, followed by the India and USA commenters. Among all the *interactional markers*, *hedges* were used most frequently by the India, UK, and USA commenters. The following *interactional marker* used by the India, UK, and USA commenters was *attitude markers*. It was observed that among all the *interactional markers*, *self-mentions* were used less by the India, UK, and USA commenters. However, the India commenters used *self-mentions* more in the sports

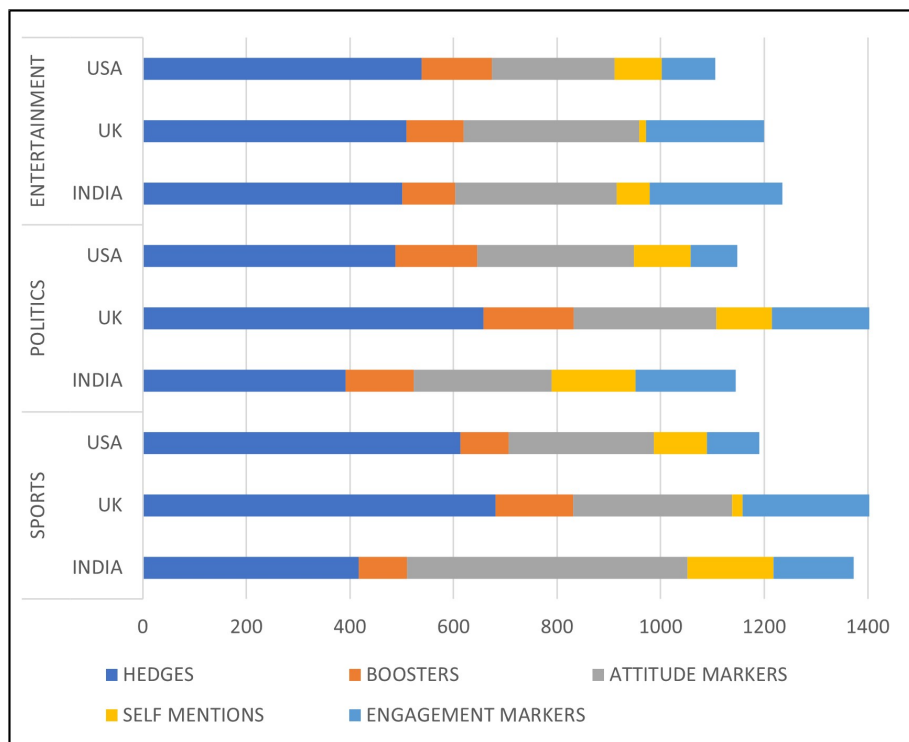


Figure 7.20: domain wise distribution (normalized) of *interactional markers* with respect to country

and politics domains in comparison to the UK and USA commenters. One possible reason for this could be that the India commenters were more inclined to proclaim authorial selves while commenting on digital newspapers.

7.7 Political Ideology Analysis of Comments

In Section 5.2.1, we have discussed the newspapers based on their political ideologies – left-wing and right-wing. Here, we provide the quantitative analysis of employing metadiscourse markers by the commenters supporting different political ideologies (in particular, left-wing and right-wing) across three different countries — the UK, India, and the USA. Similar to the domain-specific analysis of metadiscourse markers, the analysis of the use of metadiscourse markers with respect to political ideologies could be presented, taking into account different domains and countries. Considering the

fact that the domain-specific analysis of metadiscourse markers based on the political ideology is already covered in Section 7.6, we focus only on the country-specific analysis in this section.

Regarding the quantitative use of *interactive markers* with respect to political ideology, we observed that all the *interactive markers* were used by the India, UK, and USA commenters quantitatively. Table A.11 portrays that the India, UK, and USA commenters follow the order in terms of the use of *interactive markers* — *transition markers*, *frame markers*, and *code-glosses*, *endophoric markers*, and *evidentials*.

We noticed the following facts from Table A.11 and Figure 7.21:

- The left-wing commenters overall used *interactive markers* more in their comments in comparison to the right-wing commenters. One can think as a possible reason for this is that the left-wing commenters were inclined more to guide the readers and commenters through the comments.
- *Transition markers* were used most frequently by the commenters of the left-wing sub-corpora (2292.31 per 10,000 words) in comparison to the right-wing sub-corpora (1935.08 per 10,000 words).
- In the use of the *frame markers*, it was noticed that the right-wing commenters used them comparatively more (549.90 per 10,000 words) than the left-wing (407.52 per 10,000 words).
- Among the least used *interactive markers*, it was noted that *endophoric markers* and *evidentials* were used comparatively more by the left-wing than the right-wing.
- We noticed that only *code glosses* were used comparatively more by the commenters of the right-wing which indicates that the right-wing commenters were more eager to help the readers and commenters to understand the meaning of the propositions by elaborating them.

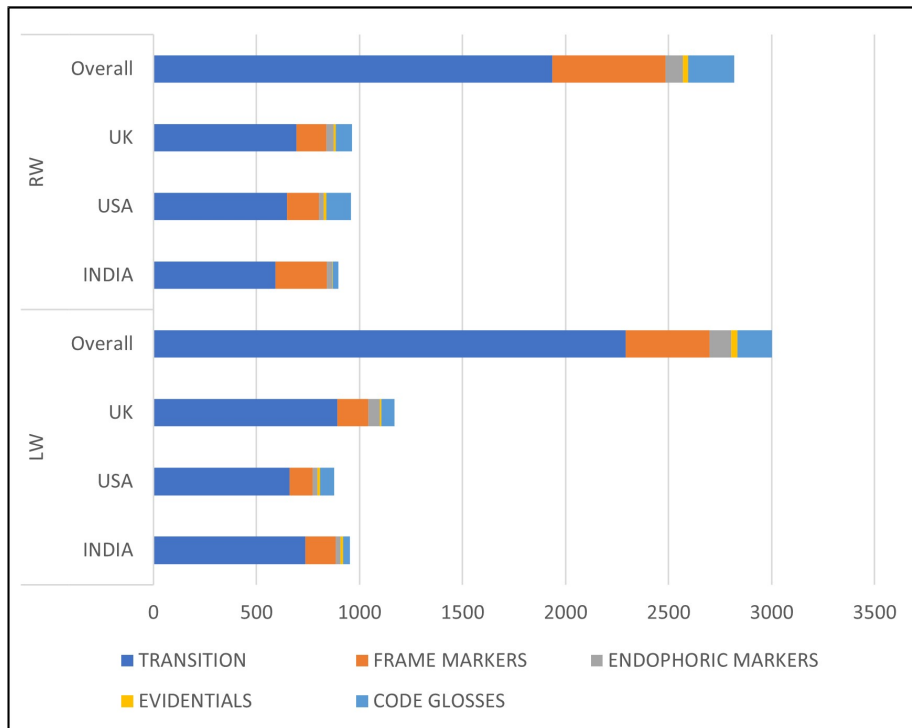


Figure 7.21: Political ideology based distribution of *interactive markers* (normalized)

With respect to the country-wise distribution of the *interactive markers* found in both the left-wing and right-wing sub-corpora, the following facts are noted from Table A.11 and Figure 7.21:

- The UK commenters of the left-wing sub-corpora used *transition markers*, *frame markers*, and *endophoric markers* more in comparison to the India and USA commenters.
- *Evidentials* and *code glosses* were used by the USA commenters more than the UK and India commenters of the left-wing sub-corpora.
- In the right-wing sub-corpora, *transition markers*, *endophoric markers* were used more by the UK commenters than the India and USA commenters.
- The India commenters used *frame markers* more than the UK and USA commenters. This suggests that the India commenters were more prone to refer to

the sequences and stages in the arguments to help the readers and commenters to understand the arguments better.

- *Evidentials* and *Code glosses* were used by the USA commenters more, followed by the UK and India commenters.

Regarding the use of *interactional markers* with respect to political ideology upholds the fact that irrespective of political ideology, all the *interactional markers* were used by the India, UK, and USA commenters. Table A.12 depicts that irrespective of political ideology, the India, USA, and UK commenters tend to use *interactional markers* in the following order: *hedges*, *attitude markers*, *engagement markers*, *boosters*, and *self-mentions*.

From Table A.12 and Figure 7.22, we noted the following facts in the use of *interactional markers*:

- Among all the *interactional markers*, *hedges* were enormously used by the commenters. The commenters of right-wing sub-corpora used *hedges* in their comments more compared to the left-wing which reveals that the right-wing commenters politely wanted to avoid direct commitment to propositions in the arguments.
- Similarly, the next marker used most frequently after *hedges* were *attitude markers*, which were used more in the comments of the right-wing sub-corpora (1528.17 per 10,000 words) than the left-wing sub-corpora (1331.51 per 10,000 words).
- The least used *interactional marker* was *self-mentions*, which was also used a bit more in the comments of the right-wing sub-corpora than the left-wing that further indicates the tendency of the right-wing commenters to uphold their authorial selves in the arguments.
- *Hedges*, *attitude markers*, and *self-mentions* were used comparatively more in the right-wing sub-corpora than the left-wing. However, in the use of *boosters*

and *engagement markers*, the left-wing commenters used them more than the right-wing.

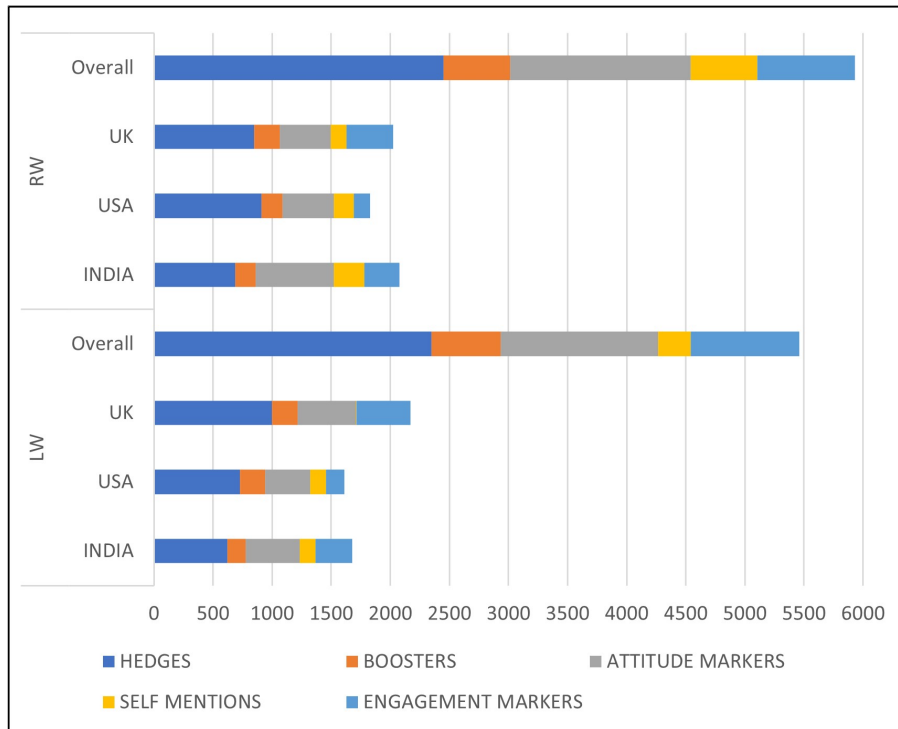


Figure 7.22: Political ideology based distribution of *interactional markers* (normalized)

Regarding the countrywise distribution of the *interactional markers*, it has been observed that most of the *interactional markers* (*hedges*, *boosters*, *attitude markers* and *engagement markers*) were used frequently by the UK commenters of left-wing sub-corpora except for *self-mentions* which were used significantly less by the UK commenters and used more by the left-wing India and USA commenters. From Table A.12 and Figure 7.22, we found the following facts on the use of *interactional markers* by the right-wing commenters:

- The USA commenters used *hedges* comparatively more than the commenters of India and the UK.
- The India commenters used *attitude markers* and *self-mentions* more in comparison to the UK and USA commenters. A possibility could be that the India

commenters were more prone to show their attitudes to propositions in front of the readers and commenters. The comparative use of *self-mentions* indicates that they willingly proclaim their authorial selves in the arguments. Also, the use of exclusive and possessive pronouns hints at the possessive nature of the India commenters.

- The *boosters* and *engagement markers* were used more by the UK commenters than the India and USA commenters which shows that the UK commenters were determined regarding the ideas and propositions and they wanted to include the readers and commenters in the arguments as well.

If we compare the quantitative use of the *interactive* and *interactional markers* by the commenters of the left-wing and right-wing sub-corpora, we can conclude that overall, the *interactional markers* were used more frequently (almost twice) than the *interactive markers*. Irrespective of political ideology, *transition markers* from *interactive markers* and *hedges* from *interactional markers* were used more frequently by the commenters. Also, irrespective of political ideology, *evidentials* from the *interactive markers* and *self-mentions* from the *interactional markers* were used comparatively less by the commenters.

7.8 Cross-cultural Analysis of Comments

The use of metadiscourse markers for specific countries is analyzed in the previous sections (cf. Section 7.2, 7.3, 7.4). This section presents the quantitative analysis of *interactive* and *interactional* markers in the context of cross-culture. By cross-culture, we mean that the commenters are from different cultural backgrounds, including national, regional, and ethnic differences. In this study, language (in our case, English) plays a crucial role as we are dealing with English news comments from three different countries. Although English is the native language for the people of the UK and USA, Indians practice English as their second language. We draw the cross-cultural

analysis of the comments based on the domains and political ideology. Considering cross-culture, the domain wise frequency and distribution of the *interactive* and *interactional* markers are portrayed in Table A.13 and Table A.14 (See in Appendix), and Figure 7.23 and Figure 7.24.

The quantitative analysis of using *interactive markers* with respect to domain reveals that all the *interactive markers* were used irrespective of domains. The commenters' frequency of using the markers follows the following order: *transition markers*, *frame markers*, *code glosses*, *endophoric markers*, and *evidentials*. The India commenters overall used *evidentials* less (12.28 per 10,000 words) in comparison to the UK (17.24 per 10,000 words) and USA commenters (25.33 per 10,000 words). Whereas the use of *interactive markers* by the India and USA commenters was almost equal, the UK commenters used *interactive markers* more in their comments. Although the UK and India commenters of their respective politics domain used *interactive markers* more compared to the other two domains (sports and entertainment), the USA commenters of the sports domain used *interactive markers* more in their comments. (See Figure 7.23).

The quantitative analysis of using *interactional markers* with respect to domains discloses the fact that all the *interactional markers* were used by the India, UK, and USA commenters. From Table A.14, it is clear that irrespective of domains and the three countries, *hedges* were used most frequently (a total of 4797.65 per 10,000 words) by the commenters (in particular, India:1309.50, USA:1639.16, UK:1848.99) among all the *interactional markers*. The rest of the *interactional markers* were used in the following order: *attitude markers* (a total of 2859.67 per 10,000 words), *engagement markers* (a total of 1751.03 per 10,000 words), *boosters* (a total of 1211.93 per 10,000 words) and *self-mentions* (a total of 836.98 per 10,000 words) by the India, USA, and UK commenters. Irrespective of domains, *hedges* were used most frequently by the UK commenters (1848.99 per 10,000 words). The India commenters used *attitude markers* (1118.25 per 10,000 words) and *self-mentions* (392.77 per 10,000 words) more in their

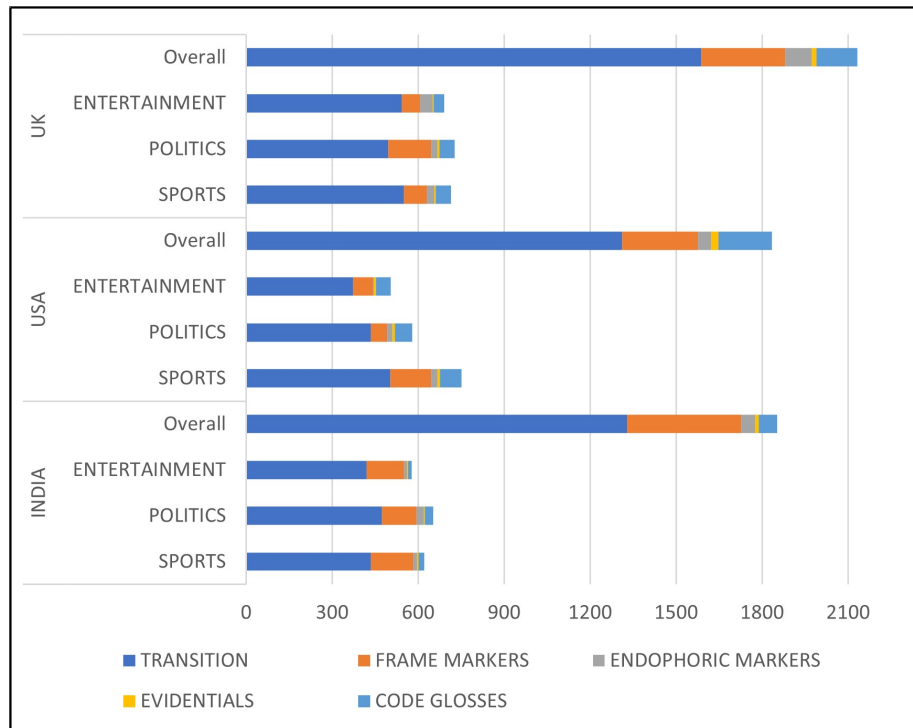


Figure 7.23: Cross-culture: domain wise frequency of *interactive markers* (normalized)

comments in comparison to the UK and USA commenters. The UK commenters used *engagement markers* and *boosters* most frequently (853.29 per 10,000 words and 434.00 per 10,000 words, respectively) compared to the India and USA commenters. Figure 7.24 depicts that the UK commenters used *intercational markers* more frequently than the India and USA commenters. It has been noted that although the India and USA commenters used *interactional markers* more in their respective sports domain, the UK commenters used them more in the politics domain.

Irrespective of domains, if we draw a comparison between the use of *interactive* and *interactional markers* based on a cross-cultural perspective, we can conclude the fact that the *interactional markers* were used more frequently by the commenters (almost twice) in comparison to the *interactive markers*.

Considering cross-culture, the frequency and distribution of *interactive* and *interactional* markers with respect to political ideology were portrayed in Figure 7.25 and Figure 7.26. for more details, please see Table A.15 and Table A.16 in the Appendix.

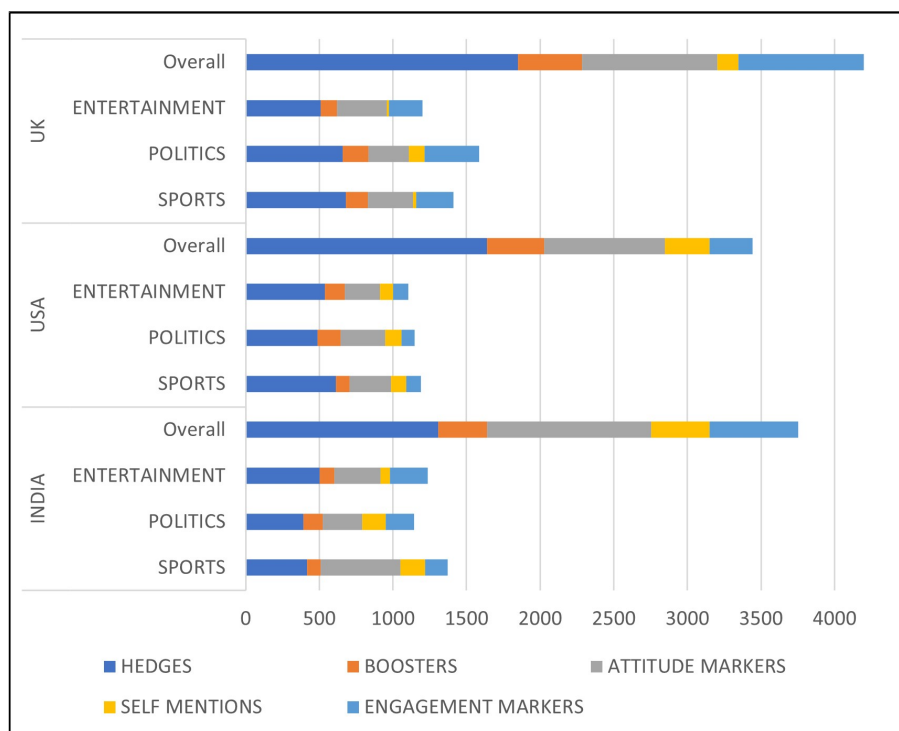


Figure 7.24: Cross-culture: domain wise frequency of *interactional markers* (normalized)

With respect to three countries, from the frequency of using *interactive markers* based on two political ideologies, we noticed the following facts:

- *Transition markers* and *endophoric markers* were most frequently used by the UK commenters, followed by the India and USA commenters.
- The India commenters used *frame markers* most (397.72 per 10,000 words) followed by the UK (293.01 per 10,000 words) and USA commenters (266.70 per 10,000 words).
- We noticed that the USA commenters used *evidentials* and *code glosses* comparatively more than the India and UK commenters. (See Figure 7.25). The comparative use of *code glosses* implies that the USA commenters have a tendency to prolong the arguments by restating the ideas and propositions. The use of *evidentials* indicates the eagerness of the USA commenters to refer to the source information of the ideas in the arguments.

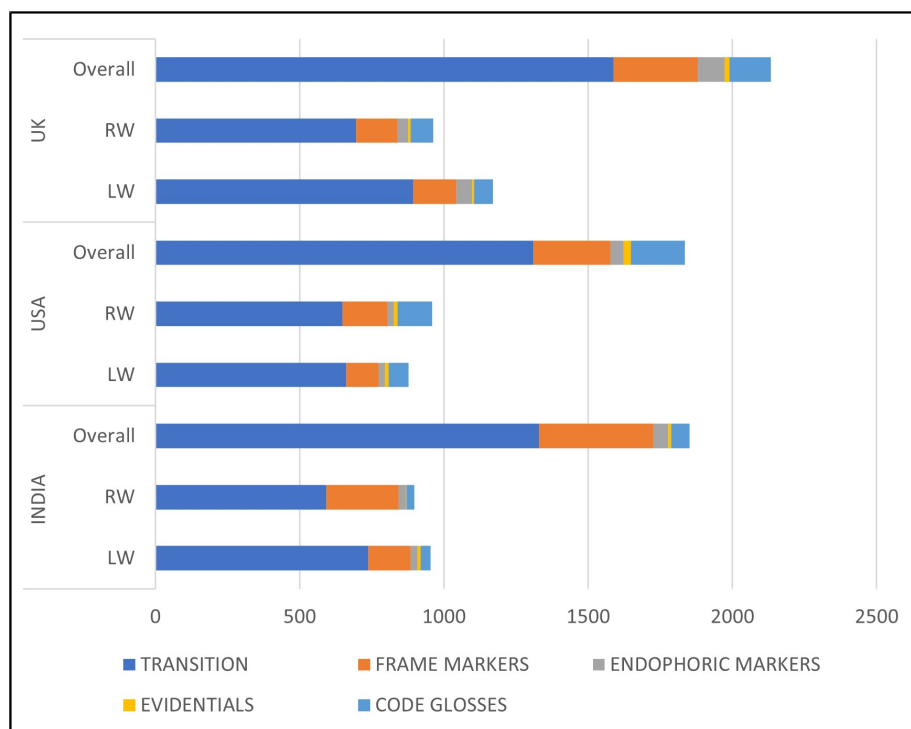


Figure 7.25: Cross-culture: political ideology based distribution of *interactive markers* (normalized)

In the use of *interactive markers* irrespective of political ideology, we observed that the left-wing India commenters used *transition markers*, *evidentials*, and *code glosses* more compared to the right-wing. Similarly, *frame markers* and *endophoric markers* were used by the India commenters of the right-wing sub-corpora more than the left-wing. The right-wing USA commenters used all the *interactive markers* more (*transition markers*, *frame markers*, *evidentials* and *code glosses*) except for *endophoric markers* which were used more (22.80 per 10,000 words) in the comments of the left-wing. We found that *transition markers*, *frame markers*, and *endophoric markers* were used more frequently by the UK commenters of the left-wing. On the contrary, the right-wing UK commenters used *evidentials* and *code glosses* more in their comments than the left-wing.

With respect to three countries, from the frequency of using *interactional markers* based on two political ideologies, we noticed the following facts:

- Among all the *interactional markers*, *hedges* were used most frequently by the commenters of three nations, particularly by the UK commenters (1848.99 per 10,000 words), followed by the USA commenters (1639.16 per 10,000 words) and the India commenters (1309.50 per 10,000 words). The enormous use of hedges by the commenters of all three nations suggests that the commenters were more likely to restrain themselves from making a full commitment in arguments.
- The subsequent *interactional marker* used most was *attitude markers* which was used most by the India commenters, followed by the UK and USA commenters.
- However, in the use of *self-mentions*, we noticed that the UK commenters used them significantly less, and the USA and India commenters used them almost 2 times more in their comments which expresses the indifference of the UK commenters in affirming authorial selves in the arguments.

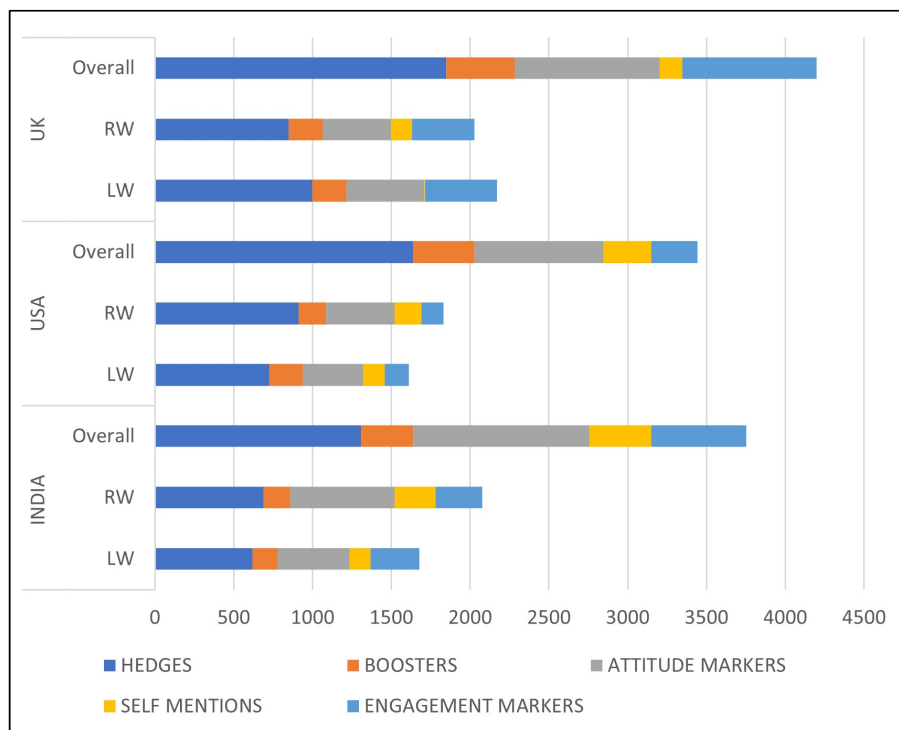


Figure 7.26: Cross-culture: political ideology based distribution of *interactional markers* (normalized)

From Figure 7.26, it was noticed that irrespective of political ideology, the right-wing India commenters used *hedges*, *boosters*, *attitude markers* and *self-mentions* more in their comments compared to the left-wing. However, only *engagement markers* were used more by the left-wing India commenters which suggests the tendency of the left-wing India commenters to build a rapport with the readers and other commenters by engaging them in the arguments. Similarly, the right-wing USA commenters used *hedges*, *attitude markers*, and *self-mentions* more than the left-wing. On the contrary, *boosters* and *engagement markers* were used more by the left-wing USA commenters than the right-wing. The left-wing UK commenters used *hedges*, *attitude markers*, and *engagement markers* more compared to the right-wing. Contrarily, the right-wing UK commenters used *boosters* and *self-mentions* comparatively more than the left-wing. This upholds the full commitment of the right-wing UK commenters to ideas and propositions in arguments and their eagerness to assert their authorial selves.

If we compare the use of *interactive* and *interactional markers*, we found that all the markers were used by the India, UK, and USA commenters. In the overall use of the markers, we found that the *interactional* markers were used comparatively more than the *interactive* ones. In using *interactive* and *interactional* markers, the UK commenters used both more than the India and USA commenters. This reflects the tendency of the UK commenters to help the readers and commenters to interpret the ideas and arguments well and also include them as participants in the arguments. It is quite surprising that despite being non-native speakers of English, the India commenters used both the *interactive* and *interactional* markers frequently after the UK. It also represents the communicative nature of the India commenters who wanted to guide the readers and commenters and also involve them in the arguments.

7.9 Conclusion

This chapter portrays the quantitative use of metadiscourse markers (*interactive* and *interactional*) in different contexts — specific to the country (India, the UK, and the USA) in Section- 7.2, 7.3, and 7.4, political ideology (left-wing and right-wing) in Section- 7.7, domain (entertainment, politics, and sports) in Section- 7.6 and cross-culture (English as native and non-native) in Section- 7.8. The distribution of the *interactive* and *interactional markers* specific to the country, political ideology, domain, and cross-culture is depicted here through figures. This chapter renders an in-depth analysis (both quantitative and qualitative) of the obtained results in Section- 7.5. Also, in this section, a detailed descriptive analysis of the occurrences is presented with examples. The analysis of the obtained results unveils the answer to our research questions that we have posed in Section 1.4. In the next chapter, we confront the research questions in detail with the help of our findings in this chapter. Also, we showed the limitations of this study and mentioned the future study that could be done based on this doctoral dissertation.

Chapter 8

Conclusions

With the rapid digital progression, nowadays, people are highly interested in digital newspapers, which are the online version of printed newspapers. The prime reason for the attraction towards these digital newspapers is that people get a scope to share their opinions in the form of comments after reading particular news. For this increasing interest of people in digital newspapers, emerging newspapers as well as well-established and prestigious newspapers, are creating an online version of their printed newspapers. This doctoral dissertation selected two categories of digital newspapers, namely, left-wing and right-wing. We analyzed the digital comments (specifically, news comments) along with the replies of the readers on the national as well as international news that were collected from the three domains (namely— politics, sports, and entertainment) and three continents (namely— European, Asian, and American).

In this last chapter, we present our concluding remarks on the thesis. Firstly, we recapitulate the major findings of this thesis to answer the research questions formulated previously. Secondly, based on the whole thesis, we assess the contribution of this research. Lastly, we comment on the limitations of this study and recommend potential avenues for future research on digital comments (specifically, news comments).

8.1 Answering to research questions

In this section, we will summarize the findings of this study by answering the research questions in the following section, which were formed in the Introduction (Section: 1.4). In our study, we chose three domains– politics, sports, and entertainment. Our selection criteria were governed by the popularity of those domains by the commenters of the UK, India, and the USA. We have discussed in detail the domain wise analysis in the usage of the metadiscourse markers by the three chosen nations in Chapter 7 (Section:7.6). The geo-location background of the commenters also affects the use of metadiscourse markers while commenting on the news for a specific domain. This implies that commenters of different geo-location backgrounds employ the metadiscourse markers not in the same manner. Therefore, we also discussed the use of metadiscourse markers based on other aspects, such as the use of English as a native and non-native language (RQ:2), different geo-location (RQ:3), and different political ideologies (RQ:4).

RQ1: Are there any differences or similarities in how writers (in our case, commenters) use metadiscourse markers across politics, sports, and entertainment domains?

In our study, we investigate the quantitative use of metadiscourse markers by the commenters while commenting on news related to different domains, particularly news related to politics, sports, and entertainment domains. In answering this research question, we would like to mention that in Section 7.6, we describe in detail the quantitative analysis of the metadiscourse markers (*interactive* and *interactional*) concerning the domains. Here we answer the question in two parts, pointing out the similarities and differences in the use of metadiscourse markers with respect to domains.

Before understanding the implication of the similarities and differences, it is crucial to know certain characteristics of the commenters belonging to politics, sports, and entertainment domains. People who comment on political news tend to be extroverts

and often more vocal. They are often careful and attentive to their claims showing much sincerity and honesty. The commenters of the sports domain are often more passionate and enthusiastic kind of people. These qualities bring up the emotional side of the commenters as they are deeply attached to the people or groups they support. The commenters of the entertainment domain are those people who abode in entertainment in order to release and distract themselves from stress. This also hints at the commenters' inclination towards escapism. Therefore, the three domains represent certain features that the commenters possess. From our corpus, with respect to domains, we noticed the following similarities in employing metadiscourse markers among commenters:

(1) Irrespective of domains, *interactional* markers were used almost 2 times more than the interactive markers by the commenters of the UK, India, and the USA. This suggests that the motive of the commenters, irrespective of the domain, was to involve other commenters and readers in the argument and build a conversation with them. (2) It was observed that the pattern of using the lexico-grammatical categories of the *interactive* and *interactional* markers was the same, irrespective of domains. It indicates that the commenters, despite belonging to different domains and possessing different characteristics, share the same thought process while commenting on particular news. (3) The distribution of the interactive and *interactional* metadiscourse markers across the domains shows that both markers were used more by the commenters of the sports domain. Following the certain characteristics of the commenters mentioned above, it can be stated that the frequent use of *interactive* and *interactional* markers by the commenters of the sports domain indicates the passionate and enthusiastic nature that encouraged them to converse with the readers and other commenters. Also, they were way more expressive with their thoughts and opinions compared to the commenters of the politics and entertainment domains. It also highlights their confidence as they guide the readers and commenters through the comments for better understanding and include them in the conversation as well. (4) In all three domains, *transition* and *frame*

markers from the *interactive* category were used more, and from the *interactional* category, *hedges* and *attitude markers* were used most frequently by the commenters. The frequent use of *transition markers* and *frame markers* by the commenters of all three nations indicates that despite belonging to different domains, they guide the readers by pointing out the semantic relation between clauses and the sequences and text stages to understand the interaction better. Also, the recurrent use of *hedges* and *attitude markers* suggests the tendency of the commenters to show their attitude and politely avoid commitment in a conversation. (5) Among the least frequent markers, irrespective of the domains, *evidentials* from the *interactive* category and *self-mentions* from the *interactional* category were used by the commenters. This implies that the commenters of the entertainment, politics, and sports domains were not so fond of proving the authenticity of a proposition by mentioning the source of information, and also, they were not so eager to show explicit reference to the author.

We draw the following differences:

(1) The first difference that we noticed was that among all the interactive markers, *code glosses* were used comparatively more in the sports domain compared to the other two domains. This suggests that the commenters of the sports domain were enthusiastic enough to guide the readers and commenters by rephrasing ideas and showing examples for better clarity. Also, this could be because sports is regarded as one of the most popular domains where commenters usually comment while being aware of the particular event where they are commenting. On the other hand, in the politics and entertainment domains, there could be a possibility that the commenters were not so eager to explain further the meaning of a particular proposition. (2) Even though *evidentials* were used significantly less among the other interactive markers across all the domains and countries, in the sports domain, they were used comparatively more than the politics and entertainment domains. This indicates that the commenters of the sports domain tend to refer to the source information more while claiming something. The comparative use of *evidentials* in the sports domain suggests that

the commenters were more eager to refer to the source of the idea compared to the other domains. Also, one possible reason for this can be that their enthusiastic nature compels them to refer to the source to make the idea more clear and more authentic to the audience. It suggests their endeavour to provide validation to the ideas and facts. However, as this doctoral dissertation deals with news comments which are informal in nature, we found that the source information is not always a reference or a citation; it may be a direct reference to a particular term and proposition. (3) Although overall *self-mentions* were used less than the other categories of *interactional* markers, in the politics domain, they were used significantly more. Following the particular characteristics of the commenters mentioned above, the frequent use of *self-mentions* in the politics domain depicts the commenters' extrovert nature in using possessive and inclusive pronouns in their comments.

In a nutshell, it can be concluded from our investigation that there are similarities as well as differences in the use of metadiscourse markers by commenters across entertainment, politics, and sports domains.

RQ2: What are the similarities or differences between English native speakers and non-native speakers who comment on digital newspapers in terms of the use of metadiscourse markers?

In our dissertation, we studied the use of metadiscourse markers by the commenters of three nations where English is the native language in the UK and the USA; however, English is the non-native and second language in India. In this context, we observed some similarities and differences in the use of metadiscourse markers by the commenters.

We found these salient similarities:

(1) Regardless of being English native and non-native speakers, overall *interactional* markers were observed to be utilized comparatively more by all the commenters. However, we found that in the UK sub-corpora, irrespective of domains, both the *in-*

teractive and *interactional* markers were used more by the commenters, followed by the India and USA sub-corpora. The fact that the UK commenters used both the *interactive* and *interactional* markers more implies that they were more likely to be more expressive with their thoughts by guiding their fellow commenters through the arguments and involving them in those arguments. (2) Irrespective of using English as their first language (i.e., native or non-native), the commenters used *transition markers* and *frame markers* more frequently among *interactive* markers, and *hedges* and *attitude markers* among *interactional* markers. The frequent use of *transition markers* and *frame markers* shows that the commenters were eager to link and sequence their arguments to maintain the flow or continuity of the comments. (3) We noticed that the least used *interactive* markers were *evidentials* for all the commenters. One possible reason could be that as digital comments (specifically, news comments) are short and informal in nature, the commenters were not bothered to show the authenticity of a particular proposition.

On the other hand, we noticed the following differences:

(1) Among the *interactive* markers, *frame markers* were used comparatively more by the non-native commenters of English (India) than the native English commenters (the UK and the USA). The significant use of *frame markers* by the non-native commenters indicates the fact that they were much more eager to indicate various functions such as sequencing, announcing ideas, shifting topics, etc., while conveying their ideas to the readers and commenters. (2) We noted that *evidentials* were used mostly by the native English commenters, and among them first was the UK and then the USA, compared to the non-native English commenters (India commenters). (3) In the use of *interactional* markers, we observed that the non-native English speakers used *attitude markers* significantly more than the native speakers of English. (4) Non-native English commenters used *self-mentions* comparatively more than native English-speaking commenters. The implication of the noted differences in (2), (3), and (4) are given in the following paragraph.

Our study provides some probable explanations of the similarities and differences in the use of metadiscourse markers by native and non-native commenters. It can be assumed that being native speakers of English, the commenters of the UK and USA would be proficient enough in using rhetorical devices while expressing their thoughts. The commenters of India, on the other hand, are non-native speakers of English, and they use English as their second language following the three-language formula, where it was asserted that in every state, three languages, Hindi, English, and a regional language, would be taught (Biswas, 2004) (See Section-2.6). Thus, the intuition was that the Indians being non-native speakers of English would be not so proficient in using rhetorical devices, particularly metadiscourse devices, while commenting on digital news, as English is not their vernacular language. However, we noticed some contradictions regarding this intuition, as the most frequent use of *self-mentions* and *frame markers* was by the India commenters. Despite being non-native speakers, the India commenters showed enough competency in using *self-mentions*, particularly in using first-person plural pronouns such as ‘we’ and ‘our’ in their comments to mention their authorial presence. As Kuo (1999) suggested, the use of personal pronouns permits the writers to emphasize their contributions. Following that, in our study, we could state that the India commenters have a tendency to emphasize their contribution by using *self-mentions* frequently. The frequent use of *frame markers* by the India commenters suggests their tendency to use short and simple sentences to mark the topic shifts, include additional information, mark the discourse stages, etc. Contrarily, the native speakers used *evidentials* comparatively more, which indicates their tendency to provide authentication and validation to the readers and commenters by referring to the source of information. The less frequent use of *evidentials* by the India commenters depicts their casual attitude. Also, they seemed to be confident enough to convince the readers and commenters about what they were stating. The frequent use of both categories by the UK commenters, followed by India and the USA, hints at the tendency to use the same pattern despite being native and non-native speakers of English.

Hence, our investigation asserts that both the native and non-native commenters were capable enough to use metadiscourse markers in an effective way.

RQ3: Do commenters from the UK, INDIA, and the USA employ metadiscourse markers similarly or differently depending on their geo-location and culture?

The answer to this research question is divided into two parts: the first will show the similarities, and the second will show the differences. In Sections 7.2, 7.3, and 7.4, we demonstrate the quantitative analysis of the metadiscourse markers employed by the commenters from India, the UK, and the USA. By geo-location, we mean the respective country from where a commenter belongs, and by culture, we mean to say the similarities and differences that we found in the cultural traits of the commenters belonging to the respective countries. Regardless of the geo-location and culture, we found some similarities and differences in the use of metadiscourse markers by the UK, India, and the USA commenters.

The similarities are as follows:

(1) In the three sub-corpora (UK, India, and USA), *interactional* markers were employed comparatively more than the *interactive* markers that point out the fact that the commenters perceived the need to involve the readers (in our case, other commenters) in the conversation. (2) Irrespective of the three nations, among the *interactive* markers, the commenters used *transition markers* and *frame markers* most frequently to help the readers and commenters to understand how they are linking and sequencing the arguments. On the other hand, among the *interactional markers*, *hedges* were used most frequently to show the commenter's partial commitments, and *attitude markers* were used most frequently to make the readers and commenters understand the attitude of the writer (in our case, the commenter who is commenting). (3) Among the three countries, the commenters most frequently used additives from *transition markers* and announcers from *frame markers*. The frequent use of the additive marker 'and' hints at the fact that the commenters were more prone to make the readers know the

linking of the arguments, and the announcer ‘will’ indicates the writer’s (in our case, the commenter who is commenting) intention to announce the following idea. (4) Irrespective of the three countries, the most frequently used lexico-grammatical category of *hedges* was modal verb, and the least used was nouns and phrases. The commenters extensively used the modal verb ‘can’, ‘would’, ‘should’, and ‘may’. However, they used the nouns ‘expectation’, ‘possibility’, ‘belief’, and the hedges-phrases ‘in theory’, ‘in general’, and ‘in part’ less frequently in the comments. Across the three countries, the commenters frequently used the attitude-adjectives ‘great’, ‘only’, and ‘good’ to indicate something superior or most suitable.

The noted differences are as follows:

(1) In the India sub-corpora, the *interactive* markers were used more in the politics domain, whereas the *interactional* markers were used more in the sports domain. On the contrary, in the UK and USA sub-corpora, both the *interactive* and *interactional* markers were used comparatively more in the politics and sports domains, respectively. One possible reason could be that the commenters of these domains were eager to guide the readers (in our case, other commenters) and, at the same time, involve them in the argument. (2) The commenters of the entertainment domain, irrespective of countries, used metadiscourse markers (*interactive & interactional*) comparatively less, which emphasizes the fact that these commenters were not so expressive, or they were not so willing to convey their views and thoughts while commenting on digital newspapers. (3) In the use of *interactive markers*, *evidentials* were used significantly more in the USA sub-corpora, followed by the UK and India sub-corpora, and *self-mentions* from the *interactional markers*, were used significantly more in the India sub-corpora, followed by the USA and the UK sub-corpora. (4) Although, overall, the use of *evidentials* was comparatively less by the UK, India, and USA commenters, among the lexico-grammatical categories of *evidentials*, the Indians used the personal marker ‘following’ more frequently, whereas the UK and USA commenters used the personal marker ‘consequence’ more. (5) Among *self-mentions*, the UK, India, and USA commenters

used ‘personal reference’ (‘we’, ‘our’), highlighting the commenters’ inclusive nature. However, the India and USA commenters did not use a single ‘self-citation’, while the UK commenters used them in a small number. This indicates the disinclination from the commenter’s side to mention his or her presence.

Our investigation supports some possible explanations of the similarities and differences in the use of metadiscourse markers (*interactive* and *interactional*) that uphold the cultural traits of the respective countries. For example, the significant use of the modal verb ‘can’ suggests that the Indians are inclined towards showing possibility or keeping the hope of something that is probable. It highlights the optimistic nature of the Indians who believe in the possibility of events. On the other hand, the excessive use of the modal-verb ‘would’ by the UK and USA commenters shows the tendency of the commenters to indicate and estimate the consequences of the events. This shows the practical nature of the UK and USA commenters who believe in the real-life world and think about the consequences beforehand. Again, the frequent use of the possessive pronouns ‘we’ and ‘our’ by the India and USA commenters reveals the fact that they are more willing to show oneness by including the readers and commenters in the conversation. However, the less frequent use of these possessive pronouns indicates the absence of this inclusive nature among the UK commenters. In short, our investigation supports that the above-mentioned characteristics that are in-built among the commenters represent their respective countries’ cultures.

RQ4: What are the similarities or differences in the use of metadiscourse markers when the commenters support a political ideology—left-wing or right-wing?

We studied the use of metadiscourse markers by the commenters supporting a particular political ideology (left-wing or right-wing) in Section-7.7. In Section-5.2.1, we described the digital newspapers used in our study that follows the left-wing and right-wing political ideologies.

The notable similarities are given below. We provide the implication of these similarities in the paragraph, which follows the differences.

(1) In the case of *interactive markers*, regardless of nations, the commenters from left-wing and right-wing sub-corpora used *transition markers* and *frame markers* more frequently.

(2) In the use of *interactional markers*, irrespective of nations, it was noticed that the commenters of both the left-wing and right-wing sub-corpora used *hedges* and *attitude markers* more in their comments.

Next, we listed the notable differences in the context of the political beliefs of the commenters. We provide the linguistic implications of these differences in the paragraph just after mentioning the differences.

(1) In our study, *evidentials* from the *interactive* metadiscourse category were used comparatively less by the commenters; however, in the left-wing sub-corpora, they were used comparatively more than the right-wing.

(2) Among the *interactive* category, *transition markers* were used comparatively more in the left-wing sub-corpora, and *frame markers* were used comparatively more in the right-wing sub-corpora.

(3) Among the *interactional* category, the less frequent marker was *self-mentions*, although they were used comparatively more in the right-wing sub-corpora than the left-wing.

(4) Among the most frequently used *interactional markers*, *hedges* and *attitude markers* were used more by the right-wing commenters.

(5) Also, *engagement markers* were seen to be used more by left-wing commenters.

If we focus on the general characteristics of left-wing and right-wing supporters, we can state that left-wing supporters tend to be much more liberal and support social equality and internationalism. Contrarily, right-wing supporters are prone to be traditional and conservative and support social hierarchy and nationalism. In our corpus, we observed that despite belonging to two different political ideologies, the

way of expression was the same for the left-wing and right-wing commenters. For example, in both the left-wing and right-wing sub-corpora, the most frequently used marker from the *interactive* category was *transition markers*, and from the *interactional* category was *hedges*. This suggests the tendency of both the left-wing and right-wing commenters to prolong the conversation along with restricting themselves from involving entirely. Also, it hints at the polite manner of avoiding any strict commitment to an idea in the conversation. However, there were some differences in the frequency of using some of the markers, for example, the frequent use of *self-mentions* in the right-wing sub-corpora. One possible reason for this could be that as traditionalism and nationalism are engraved within the hearts and minds of the right-wing supporters, it is quite likely that they would use *self-mentions* more to indicate their authorial presence. Again, the frequent use of *engagement markers* by the left-wing commenters reveals that as left-wing commenters are more liberal and believe in social equality, they are inclined to engage the readers and commenters more in the conversation by using personal pronouns, imperatives, and directives. These observations from our analysis assure that with respect to different political ideologies, there are similarities as well as differences in the use of metadiscourse markers by the commenters supporting different political ideologies.

Overall, we formed the research questions based on four factors— domains, geo-location (country) and culture, English language (native & non-native), and political ideologies. Therefore, we discussed the use of metadiscourse markers based on four aspects— different domains (RQ1), use of English as a native and non-native language (RQ2), different countries and their respective cultures (RQ3), and different political ideologies (RQ4). After analyzing our corpus, we came to the conclusion that the usage of metadiscourse markers varies, taking into account these four factors. For example, even having English as their native language, the commenters of the UK and USA (i.e., native speakers with different geographic locations and cultures) used metadiscourse markers differently across different domains and political ideologies. This reflects the

fact that cultural differences and geo-location play a vital role in their use of metadiscourse markers. Again, the commenters of the UK and the USA (native speakers with different geo-locations) used metadiscourse markers differently across different domains and political ideologies. In the case of the India commenters, domains and political ideologies play a crucial role in the use of metadiscourse markers. Hence, we may conclude from our analysis that in digital comments, the use of metadiscourse can be influenced by all of our selected factors.

8.2 Contributions of the Thesis

This work is the first work on English digital comments, particularly news comments investigating metadiscourse markers. Apart from being the pioneering work on English digital news comments, this doctoral dissertation makes the following contributions:

- This study explores rhetorical devices (especially metadiscourse devices) in English news comments. In the literature, a number of research studies have been carried out on metadiscourse markers; however, those studies were mainly based on academic discourse, which includes research articles, research abstracts, scientific articles, etc. Although a few researchers also explored metadiscourse markers in digital newspapers, their studies are restricted to newspaper articles, newspaper columns, etc. To the best of our knowledge, while writing this doctoral thesis, we have not found any research work that investigates metadiscourse markers on English news comments. Hence, this study could be considered a pioneering work that investigates the use of metadiscourse markers in digital news comments (written in English) by commenters from three nations (i.e., the UK, India, and the USA).
- When we started our dissertation, there was no corpus available for English digital news comments. Therefore, to conduct this research, we have no choice but to develop a corpus. Hence, we strongly believe that the compiled corpus that was

developed for this dissertation will represent a valuable language resource for linguistic research. We provide the details of the developed corpus in Chapter 5.

- In this dissertation, in the context of English digital news comments, for the first time, we present an in-depth quantitative as well as qualitative analysis of metadiscourse markers (*interactional* and *interactive*) (Chapter 7, Section-7.5). We probe into each *interactional* and *interactive* metadiscourse marker to show the rhetorical functions along with examples of those markers from our corpus.
- In this work, we present a detailed description of the quantitative analysis in the use of metadiscourse markers by the commenters of the three nations: India, the UK, and the USA, which belongs to three continents (Asia, Europe, and America). For each pre-selected nation, we perform the quantitative analysis of metadiscourse markers based on different domains and political ideologies. Individual analysis of the nations in the context of using metadiscourse markers could be useful to understand the expressing nature of the commenters of that particular nation. In addition, these analyses could be useful in understanding the differences as well as the similarities in their expressive nature. In Chapter 7, we provide the detailed analysis of the individual nations in Section 7.2, Section 7.3 and Section 7.4, respectively.
- This study investigates metadiscourse markers (*interactional* and *interactive*) on English news comments in the context of three different domains: entertainment, politics, and sports. In Chapter 7, Section 7.6 describes the domain wise quantitative analysis of the metadiscourse markers. This investigation gives an insight into how different domains affect the use of metadiscourse markers in the context of digital news comments across different nationalities supporting different political ideologies.
- Also, we perform a comprehensive analysis of the use of metadiscourse markers in digital news comments in the context of two well-known political ideologies,

namely– left-wing and right-wing. Section 7.7 presents the quantitative analysis of the metadiscourse markers with respect to political ideologies. We have taken into consideration the commenters from three different nations i.e., the UK, India, and the USA. We also present the analysis in using metadiscourse markers in the context of pre-selected domains as well as individual nations. Therefore, this study could play a crucial role in understanding the differences as well as similarities in using metadiscourse markers concerning two different political ideologies.

- In this dissertation, we perform an extensive analysis of the use of metadiscourse markers in digital news comments in the context of English native and non-native speakers. The commenters of the UK and USA use English as their first language. Hence, they are considered as native speakers. On the other hand, Indians are considered as non-native speakers of English; however, they use English as one of the official languages. In this context, our analysis could be of immense importance in understanding the use of metadiscourse markers by native and non-native speakers of English. This study depicts the quantitative analysis of the metadiscourse markers (*interactional* and *interactive*) taking into account English as a native and non-native language in Chapter 7, Section 7.8.

8.3 Limitations and Recommendation for Further Study

No research study is completely flawless or inclusive of all possible aspects, and our study is not an exception in this case. We tried to take into consideration all the possible aspects making this work worthwhile, however, it was not possible for us to take into account all the scenarios given the restriction in the collected data and time constraints.

The first limitation of this study is while addressing the first research question (i.e., RQ1), we restricted our scope to the three popular domains: politics, sports, and

entertainment. However, there are various other domains on digital newspapers where commenters regularly comment. We restricted our study to three domains only to meet the time constraint of this dissertation. In the future, we would like to expand this work including other interesting domains such as tourism, business, technology, etc.

Secondly, we mention in Chapter 5 (Subsection: Comment Selection) that we considered the geo-location of a commenter to be the same as the publication location of newspapers where the commenter makes a comment. For example, if a commenter comments on a particular news published in *The Times of India*, then we can assume that the commenter is a resident of India. However, it is also possible that a commenter could comment on specific news in a digital newspaper despite being from another geo-location. Also, some commenters can use fake IDs to comment on specific news. The only solution to this problem is to believe the commenters' created profiles; however, the verification of the commenters is beyond the scope of this work.

Another limitation of this study could be that we have considered collecting the news comments from three nations – the UK, India, and the USA, among which, in the UK and USA, English is the first language. However, we have not considered the other two nations, New Zealand and Australia, where people also use English as their first language. Similarly, a number of nations with English as a second language can be considered alongside India. However, to meet the time constraint, we restricted the scope of this study to three nations.

In this work, we tried for the first time to analyze metadiscourse markers on digital news comments in English. Therefore, we could not compare our findings with other existing works of the same kind. In the literature, we mentioned a lot of studies that have been carried out on metadiscourse markers specifically for academic writing, newspaper editorials, columns, etc., which are formal texts in nature. In contrast, digital news comments are short and informal text. Therefore, our work is not comparable with such studies. A few years back, [Moya Muñoz \(2016\)](#) worked on digital news comments; however, he considered only Spanish comments and followed a different model

for analysis. Hence, we could not compare our work with Moya Muñoz (2016), taking into account the incompatibility in language as well as the employed model.

Despite the contributions this study made (See Section 8.2), there is still a scope to pursue this topic further. In this dissertation, we followed the mostly adopted metadiscourse model proposed by Hyland (2005a). However, in the case of *evidentials*, we observed that as our study deals with digital comments (specifically news comments) that hold different characteristics from the academic texts, the use of *evidentials* was also different. For example, in academic texts, it is very common to cite or quote other authors or other sources of information, while in news comments, we noticed that as news comments are very short and informal in nature, the commenters are not in favour of using *evidentials* directly by citing other authors or other sources, rather they mention *evidential markers* directly within the comments. Moreover, we are highly influenced by the works of some eminent authors, such as Yang (2014), Dendale and Tasmowski (2001), Donabédian (2001), Fitneva (2001) etc., on *evidential markers* that present a different view from Hyland (2005a) on *evidentials*. In the future, in the context of metadiscourse markers, we would like to analyze digital news comments by adopting other metadiscourse models than Hyland (2005a).

Although our study attempts to explore metadiscourse markers (*interactional* and *interactive*) on English news comments, there is still a scope for us to investigate further on this topic in the future. In this study, we focused on three popular domains. So there is always a scope to explore and investigate the use of metadiscourse markers in English news comments from other domains as well, such as business, life & Style or health domain, etc. Also, we would like to include more nations where English is used as a first or second language.

In informal textual exchange, emojis are used to a great extent to express the intent and emotions of the writer. However, in formal writing, such as newspaper editorials, emojis are not used normally. Therefore, in the context of digital news comments, emojis could be taken into account as news comments are informal in nature. As a

future avenue of this research study, we would like to include emojis while analyzing the digital comments.

References

- Abdi, R. (2002), 'Interpersonal metadiscourse: An indicator of interaction and identity', *Discourse Studies* **4**(2), 139–145.
- Abdi, R., Rizi, M. T. and Tavakoli, M. (2010), 'The cooperative principle in discourse communities and genres: A framework for the use of metadiscourse', *Journal of Pragmatics* **42**(6), 1669–1679.
- Ädel, A. (2003), *The Use of Metadiscourse in Argumentative Texts by Advanced Learners and Native Speakers of English*, PhD thesis, English dept.
- Ädel, A. (2006), *Metadiscourse in L1 and L2 English*, Vol. 24, John Benjamins Publishing.
- Ädel, A. (2010), 'Just to give you kind of a map of where we are going: A taxonomy of metadiscourse in spoken and written academic English', *Nordic Journal of English Studies* **9**(2), 69–97.
- Agnihotri, R. K., Khanna, A. L. and Mukherjee, A. (1988), *Tense in Indian English: A sociolinguistic perspective*, Vol. 12, Bahri Publications.
- Akoto, O. Y. (2020), 'Metadiscourse within a discipline: A study of introduction and literature review chapters of sociology masters' theses', *Indonesian Journal of Applied Linguistics* **10**(2), 471–480.
- Albalat-Mascarell, A. and Carrió-Pastor, M. L. (2019), 'Self-representation in political

- campaign talk: A functional metadiscourse approach to self-mentions in televised presidential debates', *Journal of Pragmatics* **147**, 86–99.
- Albert, E. and Stone, J. A. (1979), *A history of English literature*, Harrap.
- Algeo, J., Lass, R. and Romaine, S. (1992), *The Cambridge history of the English language*, Vol. 3, Cambridge University Press.
- Alonso, R. A., Alonso, M. A. and Mariñas, L. T. (2012), 'Hedging: An exploratory study of pragmatic transfer in nonnative English readers' rhetorical preferences', *Ibérica* (23), 47–64.
- Alshahrani, A. A. S. (2015), 'A cross-linguistic analysis of interactive metadiscourse devices employment in native English and Arab ESL academic writings', *Theory and Practice in Language Studies* **5**(8), 1535–1542.
- Andujar, A. (2016), *Benefits of mobile instant messaging to develop ESL writing*, Vol. 62, Elsevier.
- Archer, D., Wichmann, A. and Aijmer, K. (2013), *Pragmatics: An advanced resource book for students*, Routledge.
- Babu, P. (1971), 'Prosodic features in Indian English: stress, rhythm and intonation', *Research diploma dissertation, CIE&FL, Hyderabad* **8**, 33–39.
- Baldauf, S. (2004), 'A Hindi-English jumble, spoken by 350 million', *The Christian Science Monitor* **1123**(1), 3.
- Bansal, R. K. (1976), 'The Intelligibility of Indian English. Monograph No. 4.'
- Bartlett, J. R. (1859), *Dictionary of Americanisms: a glossary of words and phrases usually regarded as peculiar to the United States*, Little, Brown,.
- Basu, D. D., Manohar, V., Banerjee, B. P. and Khan, S. A. (1999), *Introduction to the Constitution of India*, Wadhwa.

- Bergström, A. (2008), 'The reluctant audience: Online participation in the Swedish journalistic context', *Westminster Papers in Communication & Culture* 5(2).
- Biber, D., Conrad, S. and Reppen, R. (1998), *Corpus linguistics: Investigating language structure and use*, Cambridge University Press.
- Biri, Y. (2018), 'Addressing and acknowledging readers and writers. exploring metadiscourse in opinion writing online', *Półrocznik Językoznawczy Tertium* 3(1).
- Biswas, G. (2004), 'Language policy in Southeast Asia: a case study of India', *Language Policy, Planning, and Practice: A South Asian Perspective*, Oxford University Press, Oxford pp. 106–111.
- Blagojevic, S. (2004), 'Metadiscourse in academic prose: A contrastive study of academic articles written in English by English and Norwegian native speakers', *Studies about Languages* 5, 60–67.
- Blagojević, S. (2016), *Linguistic variation in research articles: When discipline tells only part of the story*, Bethany Gray, John Benjamins Publishing Company, Amsterdam/Philadelphia (2015), p. 222, ISBN 9789027203793, EUR 95.00, USD 143.00, Elsevier.
- Blake, N. (1996), *A history of the English language*, Macmillan International Higher Education.
- Bolton, K. (2012a), 'Varieties of world Englishes', *The encyclopedia of applied linguistics*.
- Bolton, K. (2012b), World Englishes and Asian Englishes: A survey of the field, in 'English as an international language in Asia: Implications for language education', Springer, pp. 13–26.
- Bondi, M. (2010), 'Metadiscursive practices in introductions. phraseology and semantic sequences across genres', *NJES* 9, 99–123.

- Boshrabadi, A. M., Biria, R. and Zavari, Z. (2014), 'A cross cultural analysis of textual and interpersonal metadiscourse markers: The case of economic articles in english and persian newspapers', *Advances in Language and Literary Studies* **5**(2), 59–66.
- Caers, R., De Feyter, T., De Couck, M., Stough, T., Vigna, C. and Du Bois, C. (2013), 'Facebook: A literature review', *New media & society* **15**(6), 982–1002.
- Cao, F. and Hu, G. (2014), 'Interactive metadiscourse in research articles: A comparative study of paradigmatic and disciplinary influences', *Journal of Pragmatics* **66**, 15–31.
- Carrió-Pastor, M. L. (2012), 'A contrastive analysis of epistemic modality in scientific English', *Revista de Lengua para fines específicos* **18**, 115–132.
- Carrió-Pastor, M. L. (2019a), 'Do writers express the same attitude in historical genres?', *Writing history in Late Modern English. Explorations of the Coruña Corpus* **12**, 237–259.
- Carrió-Pastor, M. L. (2019b), 'Phraseology in specialised language: A contrastive analysis of mitigation in academic papers', in 'International Conference on Computational and Corpus-Based Phraseology', Springer, pp. 61–72.
- Carrió-Pastor, M. L. (2020), 'Epistemic modals in academic English', *Re-Assessing Modalising Expressions: Categories, co-text, and context* Edited by Pascal Hohaus and Rainer Schulze, pp. 253–280.
- Carrió-Pastor, M. L. and Alonso-Almeida, F. (2019), 'Constructing legitimation in Scottish newspapers: The case of the independence referendum', *Discourse Studies* **21**(6), 621–635.
- Carrió-Pastor, M. L. and Muñiz Calderón, R. (2015), 'A contrastive analysis of metadiscourse features in business e-mails written by non-native speakers of english', *Procedia Social and Behavioral Sciences* **173**(1), 214–221.

- Chaudhary, C. (1989), *Some aspects of the phonology of Indian English*, Ranchi, India: Jayaswal.
- Cheng, X. and Steffensen, M. S. (1996), 'Metadiscourse: A technique for improving student writing', *Research in the Teaching of English* pp. 149–181.
- Crismore, A. (1983), *Metadiscourse: What it is and how it is used in school and non-school social science texts*, Champaign, Ill.: University of Illinois at Urbana-Champaign.
- Crismore, A. and Farnsworth, R. (1990), *Metadiscourse in popular and professional science discourse*, Vol. 118, *The writing scholar: Studies in academic discourse*.
- Crismore, A., Markkanen, R. and Steffensen, M. S. (1993), 'Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students', *Written communication* **10**(1), 39–71.
- Crompton, P. (2012), 'Characterising hedging in undergraduate essays by Middle Eastern students', *The Asian ESP Journal* **8**(2), 55–78.
- Crystal, D. (1987), *The Cambridge encyclopedia of language*, Vol. 1, Cambridge University Press.
- Crystal, D. (1988), *The English Language*, London: Penguin Books. Ltd.
- Crystal, D. (1997), *English as a global language*, Cambridge University Press.
- Crystal, D. (2011), *A dictionary of linguistics and phonetics*, Vol. 30, John Wiley & Sons.
- Dafouz-Milne, E. (2008), 'The pragmatic role of textual and interpersonal metadiscourse markers in the construction and attainment of persuasion: A cross-linguistic study of newspaper discourse', *Journal of Pragmatics* **40**(1), 95–113.

- Dargahi Nobari, A., Reshadatmand, N. and Neshati, M. (2017), Analysis of telegram, an instant messaging service, *in* 'Proceedings of the 2017 ACM on Conference on Information and Knowledge Management', pp. 2035–2038.
- Dash, N. S. (2008), *Corpus linguistics: An introduction*, Pearson Education India.
- Dehghan, M. and Chalak, A. (2016), 'Code glosses in academic writing: The comparison of Iranian and native authors', *Research in English Language Pedagogy* **3**(2), 21–29.
- Delbridge, A. and Bernard, J. R. (1981), *The Macquarie Dictionary*, Macquarie Library Sydney.
- Dendale, P. and Tasmowski, L. (2001), 'Introduction: Evidentiality and related notions', *Journal of pragmatics* **33**(3), 339–348.
- Desrul, D. R. K. and Romadhony, A. (2019), Abusive language detection on Indonesian online news comments, *in* '2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)', IEEE, pp. 320–325.
- Diakopoulos, N. and Naaman, M. (2011a), Topicality, time, and sentiment in online news comments, *in* 'CHI'11 Extended Abstracts on Human Factors in Computing Systems', pp. 1405–1410.
- Diakopoulos, N. and Naaman, M. (2011b), Towards quality discourse in online news comments, *in* 'Proceedings of the ACM 2011 conference on Computer supported cooperative work', pp. 133–142.
- Dillard, J. L. (2014), *A history of American English*, Routledge.
- Donabédian, A. (2001), 'Towards a semasiological account of evidentials: An enunciative approach of-er in Modern Western Armenian', *Journal of pragmatics* **33**(3), 421–442.

- Douai, A. and Nofal, H. K. (2012), 'Commenting in the online Arab public sphere: Debating the Swiss minaret ban and the "Ground Zero Mosque" online', *Journal of Computer-Mediated Communication* **17**(3), 266–282.
- D'souza, J. (2001), 'Contextualizing range and depth in Indian English', *World Englishes* **20**(2), 145–160.
- Dueñas, P. M. (2007), 'I/we focus on...': A cross-cultural analysis of self-mentions in business management research articles', *Journal of English for Academic Purposes* **6**(2), 143–162.
- Dueñas, P. M. (2010), 'Attitude markers in business management research articles: A cross-cultural corpus-driven approach', *International Journal of Applied Linguistics* **20**(1), 50–72.
- Fairclough, N. (2003), *Analysing discourse: Textual analysis for social research*, Psychology Press.
- Fauzi, I. and Angkasawati, P. (2019), 'The use of listening logs through whatsapp in improving listening comprehension of efl students', *JOALL (Journal of Applied Linguistics and Literature)* **4**(1), 13–26.
- Finch, G. (2000), *Linguistic terms and concepts*, Springer.
- Fishman, J. A. (1977), 'English in the context of international societal bilingualism', *The spread of English: the sociology of English as an additional language* pp. 329–36.
- Fitneva, S. A. (2001), 'Epistemic marking and reliability judgments: Evidence from Bulgarian', *Journal of pragmatics* **33**(3), 401–420.
- Foucault, M. (1972), 'The archaeology of knowledge and the discourse on language, trans', *AM Sheridan Smith (New York: Pantheon, 1972)* **129**(17), 130–131.
- Friginal, E. and Hardy, J. (2013), *Corpus-based sociolinguistics: A guide for students*, Routledge.

- Fuertes-Olivera, P. A., Velasco-Sacristán, M., Arribas-Baño, A. and Samaniego-Fernández, E. (2001), 'Persuasion and advertising English: Metadiscourse in slogans and headlines', *Journal of pragmatics* **33**(8), 1291–1307.
- Gargesh, R. (2008), 'Indian English: Phonology', *Varieties of English* **4**, 231–243.
- Gee, J. P. (2004), *An introduction to discourse analysis: Theory and method*, Routledge.
- Gerhards, J. and Schäfer, M. S. (2010), 'Is the internet a better public sphere? Comparing old and new media in the USA and Germany', *New media & society* **12**(1), 143–160.
- GezegİN, B. B. and Melike, B. (2020), 'Metadiscourse in academic writing: A comparison of research articles and book reviews', *Eurasian Journal of Applied Linguistics* **6**(1), 45–62.
- Gillaerts, P. and Van de Velde, F. (2010), 'Interactional metadiscourse in research article abstracts', *Journal of English for Academic purposes* **9**(2), 128–139.
- Gjesdahl, A. (2008), *News as a Cultural Product: A cross-cultural Study of Language Use in British and Norwegian Online Newspapers*, Master's thesis, The University of Bergen.
- Gomez, P. (2009), *British and American English pronunciation differences*, Cambridge: Mayflower Press.
- Gon, S. and Rawekar, A. (2017), 'Effectivity of e-learning through whatsapp as a teaching learning tool', *MVP Journal of Medical Sciences* **4**(1), 19–25.
- Görlach, M. (1988), 'English as a world language—the state of the art', *English World-Wide* **9**(1), 1–32.
- Graddol, D. (1997), 'The Future of English? London', *The British Council*.

- Gupta, R., Abbi, A. and Aggarwal, K. S. (1995), *Language and the state: Perspectives on the eighth schedule*, Creative Books.
- Halliday, M. A. K. et al. (1964), 'The linguistic sciences and language teaching'.
- Hazaea, A. N. and Alzubi, A. A. (2016), 'The Effectiveness of Using Mobile on EFL Learners' Reading Practices in Najran University.', *English language teaching* **9**(5), 8–21.
- Hille, S. and Bakker, P. (2014), 'Engaging the social news user: Comments on news sites and Facebook', *Journalism Practice* **8**(5), 563–572.
- Hooi, C. M., Tan, H., Lee, G. I. and Danarajan, S. S. V. (2020), 'Texts with metadiscourse features are more engaging: A fact or a myth?', *3L: Language, Linguistics, Literature* **26**(4).
- Huang, Y., Wang, H. and Tang, J. (2020), 'A Study of Interactional Metadiscourse in English Travel Blogs', *Open Journal of Modern Linguistics* **10**(6), 785–793.
- Hyland, K. (1995), 'The author in the text: Hedging scientific writing.', *Hong Kong papers in linguistics and language teaching* **18**, 33–42.
- Hyland, K. (1996), 'Talking to the academy: Forms of hedging in science research articles', *Written communication* **13**(2), 251–281.
- Hyland, K. (1998a), 'Boosting, hedging and the negotiation of academic knowledge', *Text & Talk* **18**(3), 349–382.
- Hyland, K. (1998b), 'Persuasion and context: The pragmatics of academic metadiscourse', *Journal of pragmatics* **30**(4), 437–455.
- Hyland, K. (2001), 'Humble servants of the discipline? self-mention in research articles', *English for specific purposes* **20**(3), 207–226.

- Hyland, K. (2004), 'Disciplinary interactions: Metadiscourse in l2 postgraduate writing', *Journal of second language writing* **13**(2), 133–151.
- Hyland, K. (2005a), *Metadiscourse: Exploring interaction in writing*, Bloomsbury Publishing.
- Hyland, K. (2005b), 'Stance and engagement: A model of interaction in academic discourse', *Discourse studies* **7**(2), 173–192.
- Hyland, K. (2007), 'Applying a gloss: Exemplifying and reformulating in academic discourse', *Applied linguistics* **28**(2), 266–285.
- Hyland, K. (2010), 'Metadiscourse: Mapping interactions in academic writing', *Nordic Journal of English Studies* **9**(2), 125–143.
- Hyland, K. and Jiang, F. K. (2018), "'In this paper we suggest": Changing patterns of disciplinary metadiscourse', *English for Specific Purposes* **51**, 18–30.
- Hyland, K. and Jiang, F. K. (2020), 'Text-organizing metadiscourse: Tracking changes in rhetorical persuasion', *Journal of Historical Pragmatics* **21**(1), 137–164.
- Hyland, K. and Jiang, F. K. (2022), Metadiscourse: the evolution of an approach to texts, in 'Text & Talk', De Gruyter Mouton, pp. 31–44.
- Hyland, K. and Tse, P. (2004), 'Metadiscourse in academic writing: A reappraisal', *Applied linguistics* **25**(2), 156–177.
- Hyland, K. and Zou, H. J. (2020), 'In the frame: Signalling structure in academic articles and blogs', *Journal of Pragmatics* **165**, 31–44.
- Ihlström Eriksson, C. (2004), The evolution of a new (s) genre, PhD thesis, Department of Informatics, Gothenburg University.
- Intaraprawat, P. and Steffensen, M. S. (1995), 'The use of metadiscourse in good and poor esl essays', *Journal of second language writing* **4**(3), 253–272.

- Jalilifar, A. and Alipour, M. (2007), 'How explicit instruction makes a difference: Metadiscourse markers and efl learners' reading comprehension skill', *Journal of College Reading and Learning* **38**(1), 35–52.
- Jones, R. H., Chik, A. and Hafner, C. A. (2015), *Discourse and digital practices: Doing discourse analysis in the digital age*, Routledge.
- Kachru, B. (1983), *The Indianization of English: The English Language in India*, Delhi. Oxford University Press.
- Kachru, B. (1994), 'English in South Asia', *The Cambridge* .
- Kachru, B. B. (1986), 'The Indianization of English', *English Today* **2**(2).
- Kachru, B. B. (1990), 'World englishes and applied linguistics', *World Englishes* **9**(1), 3–20.
- Kachru, B. B. (1998), 'English as an Asian language', *Links & letters* pp. 89–108.
- Kachru, B. B., Kachru, Y. and Nelson, C. L. (2009), *The handbook of world Englishes*, Vol. 48, John Wiley & Sons.
- Kachru, B. B., Quirk, R. and Widdowson, H. (1985), 'Standards, codification and sociolinguistic realism', *World Englishes. Critical Concepts in Linguistics* pp. 241–270.
- Kalogeropoulos, A., Negredo, S., Picone, I. and Nielsen, R. K. (2017), 'Who shares and comments on news?: A cross-national comparative analysis of online and social media participation', *Social media+ society* **3**(4), 2056305117735754.
- Kaya, F. and Sofu, H. (2020), 'Exploring effects of explicit teaching of metadiscourse markers on efl students' writing proficiency', *The Reading Matrix: An International Online Journal* **20**(2).

- Keene, D. (2000), 'Metropolitan values: migration, mobility and cultural norms, London 1100–1700', *L. Wright (Ed.) 2000*, 93–114.
- Khopitak, A. (2015), *Metadiscourse In Newspaper Opinion Articles: A Comparative Study Between American And Thai English-language Newspapers*, PhD thesis, Thammasat University.
- Kirkpatrick, A. (2010), *The Routledge handbook of world Englishes*, Routledge.
- Knight, D., Adolphs, S. and Carter, R. (2012), 'Canelc: constructing an e-language corpus', *Corpora* **9**(1), 29–56.
- Knight, D., Adolphs, S. and Carter, R. (2013), 'Formality in digital discourse: a study of hedging in canelc', in 'Yearbook of Corpus Linguistics and Pragmatics 2013', Springer, pp. 131–152.
- Kolhatkar, V. and Taboada, M. (2017), 'Constructive language in news comments', in 'Proceedings of the first workshop on abusive language online', pp. 11–17.
- Kopple, W. J. V. (1985), 'Some exploratory discourse on metadiscourse', *College composition and communication* pp. 82–93.
- Korau, S. M. and Aliyu, M. M. (2020), 'Use of Metadiscourse in the Persuasive Writing of Nigerian Undergraduates.', *English Language Teaching* **13**(4), 104–113.
- Kothari, A., Magdy, W., Darwish, K., Mourad, A. and Taei, A. (2013), 'Detecting comments on news articles in microblogs', in 'Proceedings of the International AAAI Conference on Web and Social Media', Vol. 7, pp. 293–302.
- Kovecses, Z. (2000), *American English: an introduction*, Broadview Press.
- Krishnaswami, N. and Sriraman, T. (1995), 'English teaching in India: past, present and future', *English Language Teaching in India: Issues and innovations London: Sage Publications*.

- Kuhi, D. and Mojood, M. (2014), 'Metadiscourse in newspaper genre: A cross-linguistic study of English and Persian editorials', *Procedia-Social and Behavioral Sciences* **98**, 1046–1055.
- Kumpf, E. P. (2000), 'Visual metadiscourse: Designing the considerate text', *Technical Communication Quarterly* **9**(4), 401–424.
- Kuo, C.-H. (1999), 'The use of personal pronouns: Role relationships in scientific journal articles', *English for specific purposes* **18**(2), 121–138.
- Kwak, H., Lee, C., Park, H. and Moon, S. (2010), What is twitter, a social network or a news media?, in 'Proceedings of the 19th international conference on World wide web', pp. 591–600.
- Lee, C. K. (2011), Micro-blogging and status updates on facebook: Texts and practices, in 'Digital discourse: Language in the new media', Oxford University Press Oxford, pp. 110–128.
- Liao, J. (2020), Metadiscourse, cohesion, and engagement in 12 written discourse, in 'Languages', Vol. 5, Multidisciplinary Digital Publishing Institute, p. 25.
- Lüdeling, A. and Kytö, M. (2008), *Corpus linguistics: An international handbook*, Walter de Gruyter GmbH.
- Lupton, D. (1992), 'Discourse analysis: A new methodology for understanding the ideologies of health and illness', *Australian Journal of Public Health* **16**(2), 145–150.
- MacNeil, R., McCrum, R. and Cran, W. (1986), *The story of English*, Barnes & Noble.
- Mameghani, A. A. and Ebrahimi, S. F. (2017), 'Realization of attitude and engagement markers in students' presentations', *International Journal of Applied Linguistics and English Literature* **6**(2), 73–77.
- Mardani, T. (2017), 'Metadiscourse markers: A contrastive study of translated and non-translated persuasive texts', *Language and Translation* **7**(2), 73–79.

- Masica, C. (1972), *The Sound System of Indian English. Monograph No. 7.*, ERIC, Hyderabad, India.
- Mauranen, A. (1993), 'Contrastive ESP rhetoric: Metatext in Finnish-English economics texts', *English for specific Purposes* **12**(1), 3–22.
- McArthur, T. (1987), 'The English languages?', *English Today* **3**(3), 9–13.
- McArthur, T. (2002), *The Oxford guide to world English*, Citeseer.
- McClosky, H. and Chong, D. (1985), 'Similarities and differences between left-wing and right-wing radicals', *British Journal of Political Science* **15**(3), 329–363.
- Mehrotra, R. R. and Mugler, F. (2000), 'Indian english. texts and interpretation-(1998). varieties of english around the world, text series vol. 7. amsterdam/philadelphia: John benjamins publishing company.', *South Pacific Journal of Psychology* **12**, 64–66.
- Meyer, B. J. (1975), *The organization of prose and its effects on memory*, Vol. 1, North-Holland.
- Meyer, C. F. (2002), *English corpus linguistics: An introduction*, Cambridge University Press.
- MIB (2001), 'India 2001: A reference annual', *Ministry of Information and Broadcasting, Government of India [MIB]*.
- Milne, E. D. (2003), 'Metadiscourse revisited: a contrastive study of persuasive writing in professional discourse. regreso al metadiscurso: estudio contrastivo de la persuasión en el discurso profesional', *Estudios ingleses de la Universidad Complutense* **11**, 29–52.
- Moya Muñoz, P. (2016), Análisis comparativo de las estrategias metadiscursivas en los comentarios de noticias en la prensa digital española y chilena, PhD thesis, Universitat Politècnica de València.

- Mukherjee, J. (2007), 'Steady states in the evolution of New Englishes: Present-day Indian English as an equilibrium', *Journal of English Linguistics* **35**(2), 157–187.
- Mur-Dueñas, P. (2011), 'An intercultural analysis of metadiscourse features in research articles written in English and in Spanish', *Journal of pragmatics* **43**(12), 3068–3079.
- Nehru, J. (2008), *Discovery of India*, Penguin UK.
- Nielsen, R. K. (2015), *Local journalism: The decline of newspapers and the rise of digital media*, IB Tauris.
- Nilsson, B. and Carlsson, E. (2014), 'Swedish politicians and new media: Democracy, identity and populism in a digital discourse', *New Media & Society* **16**(4), 655–671.
- Noble, W. (2010), 'Understanding metadiscoursal use: Lessons from a 'local' corpus of learner academic writing', *Nordic Journal of English Studies* **9**(2), 145–169.
- Noorian, M. and Biria, R. (2017), 'Interpersonal metadiscourse in persuasive journalism: A study of texts by American and Iranian EFL columnists', *Journal of Modern Languages* **20**(1), 64–79.
- Pandey, P. K. (1994), 'On a description of the phonology of Indian English', *R. K. Agnihotri et L. Khanna (ed)* .
- Parasher, S. (1983), 'Indian English: Certain grammatical, lexical and stylistic features', *English World-Wide* **4**(1), 27–42.
- Park, D., Sachar, S., Diakopoulos, N. and Elmqvist, N. (2016), Supporting comment moderators in identifying high quality online news comments, *in* 'Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems', pp. 1114–1125.
- Peacock, M. (2006), 'A cross-disciplinary comparison of boosting in research articles', *Corpora* **1**(1), 61–84.
- Pedersen, O. K. (2009), 'Discourse analysis', *Preuzeto Sa* **20**, 2009–65.

- Pennycook, A. (1994), *The cultural politics of English as an international language*, Taylor & Francis.
- Pérez-Llantada, C. (2010), 'The discourse functions of metadiscourse in published academic writing: issues of culture and language', *Nordic Journal of English Studies* **9**(2), 41–68.
- Phillipson, R. (1992), *Linguistic Imperialism*, Oxford University Press.
- Pickering, J. (1816), *A Vocabulary; Or, Collection of Words and Phrases: Which Have Been Supposed to be Peculiar to the United States of America. To which is Prefixed an Essay on the Present State of the English Language in the United States.*, Cummings and Hilliard, No. 1 Cornhill.
- Pingali, S. (2009), *Indian English*, Edinburgh University Press.
- Pinto-Coelho, Z., Carvalho, A. and Castro Seixas, E. (2019), 'News discourse and readers' comments: expanding the range of citizenship positions?', *Journalism* **20**(6), 733–751.
- Pooresfahani, A. F., Khajavy, G. H. and Vahidnia, F. (2012), 'A contrastive study of metadiscourse elements in research articles written by Iranian applied linguistics and engineering writers in English', *English Linguistics Research* **1**(1), 88.
- Rad, M. R. (2020), 'Iranian EFL Learners Use of Metadiscourse Markers in Argumentative Writing', *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal* **3**(3), 1353–1362.
- Rahimpour, S. and Faghieh, E. (2009), 'Contrastive rhetoric of English and Persian written texts: Metadiscourse in applied linguistics research articles', *Rice Working Papers in Linguistics* **1**.
- Rao, G. S. (1954), *Indian words in English: a study in Indo-British cultural and linguistic relations*, Clarendon Press.

- Sailaja, P. (2012), 'Indian English: Features and sociolinguistic aspects', *Language and Linguistics Compass* **6**(6), 359–370.
- Salek, M. (2014), 'A diagram of interactive and interactional markers in different parts of English research articles', *Journal of Language Sciences & Linguistics* **2**(3), 55–66.
- Schiffrin, D. (1980), 'Meta-talk: Organizational and evaluative brackets in discourse', *Sociological inquiry* **50**(3-4), 199–236.
- Schneider, E. W. (1997), 'Chaos theory as a model for dialect variability and change', *Issues and methods in dialectology* pp. 22–36.
- Schneider, E. W. (2003), 'The dynamics of New Englishes: From identity construction to dialect birth', *Language* **79**(2), 233–281.
- Sheorey, R. and Nayar, P. B. (2002), 'Learning and teaching English in India: Looking in from outside', *Indian journal of applied linguistics* **28**(2).
- Siddique, A. R., Mahmood, M., Akhter, N. and Arslan, F. (2018), 'Hedges as metadiscourse in Pakistani English newspaper editorials: A corpus based study', *International Journal of Development Research* **8**(5), 20478–20486.
- Skorczyńska, H. and Carrió-Pastor, M. L. (2021), 'A cross-disciplinary study of verb boosters in research articles from engineering, medicine and linguistics: Frequency and co-text variations', *Revista signos* **54**(106), 575–599.
- Smith, L. E. (1976), 'English as an international auxiliary language', *RELC journal* **7**(2), 38–42.
- Sperber, D. and Wilson, D. (2002), 'Pragmatics, modularity and mind-reading', *Mind & Language* **17**(1-2), 3–23.
- Sridhar, S. (2013), The world of the multilingual: An indian perspective, in 'Keynote address at the Linguistic Society of India Annual Conference and first international conference, Central Institute of Indian Languages, Mysore'.

- Sridhar, S. (2020), 'Indian english', *The Handbook of Asian Englishes* pp. 241–277.
- Suau Martínez, J. et al. (2015), Citizens and online media participation. Attitudes and motivations towards participatory journalism and other online practices in London and Barcelona, PhD thesis, Universitat Ramon Llull.
- Sultan, A. H. (2011), 'A contrastive study of metadiscourse in English and Arabic linguistics research articles', *Acta Linguistica* **5**(1), 28.
- Sutikno, T., Handayani, L., Stiawan, D., Riyadi, M. A. and Subroto, I. M. I. (2016), 'Whatsapp, viber and telegram: Which is the best for instant messaging?', *International Journal of Electrical & Computer Engineering (2088-8708)* **6**(3).
- Swales, J. (1990), *Genre analysis: English in academic and research settings*, Cambridge University Press.
- Thompson, G. (2001), 'Interaction in academic writing: Learning to argue with the reader', *Applied linguistics* **22**(1), 58–78.
- Thurlow, C. (2018), 'Digital discourse: Locating language in new/social media', *The SAGE handbook of social media. New York: Sage* pp. 135–145.
- Trudgill, P. and Hannah, J. (2008), *International English: A guide to varieties of Standard English*, Routledge.
- Tsagkias, M., Weerkamp, W. and De Rijke, M. (2009), Predicting the volume of comments on online news stories, in 'Proceedings of the 18th ACM conference on Information and knowledge management', ACM, pp. 1765–1768.
- Tully, M. (1997), 'English: an advantage to India?', *ELT Journal* **51**(2), 157–164.
- Vahid Dastjerdi, H. et al. (2012), 'The impact of explicit instruction of metadiscourse markers on efl learners' writing performance', *Journal of Teaching Language Skills* **29**(2), 155–174.

- Valero-Garcés, C. (1996), 'Contrastive ESP rhetoric: Metatext in Spanish-English economics texts', *English for Specific Purposes* **15**(4), 279–294.
- Walková, M. (2019), 'A three-dimensional model of personal self-mention in research papers', *English for Specific Purposes* **53**, 60–73.
- Wang, L., Zhang, Y., WANG, L. and ZHANG, Y. (2016), 'A study of metadiscourse features in English news reports', *International Journal of Arts and Commerce* **5**(6), 75–83.
- Ward, A. W. and Waller, A. R. (1916), *The Cambridge History of English Literature: The nineteenth century. II*, Vol. 13, Cambridge University Press.
- Weber, P. (2014), 'Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments', *New Media & Society* **16**(6), 941–957.
- Webster, N. (2018), *Dissertations on the English Language: Tales by Musaeus, Tieck, Richter*, BoD–Books on Demand.
- Wharton, T. (2003), 'Natural pragmatics and natural codes', *Mind & language* **18**(5), 447–477.
- Whitworth, G. C. (1885), *An Anglo-Indian dictionary: a glossary of Indian terms used in English, and of such English or other non-Indian terms as have obtained special meanings in India*, London: K. Paul, Trench.
- Williams, J. M. (1981), Literary style: The personal voice, in 'Style and variables in English', Cambridge, MA: Winthrop, pp. 117–216.
- Williams, J. M. and Nadel, I. B. (1989), *Style: Ten lessons in clarity and grace*, Scott, Foresman Glenview, IL.
- Wilson, D. (2005), 'New directions for research on pragmatics and modularity', *Trends in Linguistics Studies and Monographs* **161**, 375.

- Wiltshire, C. R. and Harnsberger, J. D. (2006), 'The influence of Gujarati and Tamil L1s on Indian English: A preliminary study', *World Englishes* **25**(1), 91–104.
- Yang, L. (2014), 'Evidentiality in English Research Articles of Applied Linguistics: From the Perspective of Metadiscourse.', *Journal of Language Teaching & Research* **5**(3).
- Yao, G. (2022), Metadiscourse use in Spanish academic writing: exploring the interface of nativeness and expertise, in 'Proyecto de investigación:', Universidad de Murcia.
- Yazdani, S., Sharifi, S. and Elyassi, M. (2014), 'Interactional Metadiscourse in English and Persian News Articles about 9/11', *Theory and Practice in Language Studies*, *4* (2), 428-434 .
- Yule, H. and Burnell, A. C. (1903), *Hobson-Jobson: a glossary of Anglo-Indian words and phrases*, London: John Murray. New edition.
- Zannettou, S., ElSherief, M., Belding, E., Nilizadeh, S. and Stringhini, G. (2020), Measuring and characterizing hate speech on news websites, in '12th ACM Conference on Web Science', pp. 125–134.
- Zarei, G. R. and Mansoori, S. (2011), 'A Contrastive Study on Metadiscourse Elements Used in Humanities vs. Non Humanities across Persian and English.', *English Language Teaching* **4**(1), 42–50.
- Zhang, B. and Jiang, Z. (2008), 'On grammatical differences between daily British and American English', *Asian Social Science* **4**(6), 69–73.
- Ziegele, M., Breiner, T. and Quiring, O. (2014), 'What creates interactivity in online news discussions? An exploratory analysis of discussion factors in user comments on news items', *Journal of Communication* **64**(6), 1111–1138.

Appendix A

A.1 Quantitative Analysis of India Comments

Table A.1: India: frequency of *interactive* markers (normalized results)

	SPORTS			POLITICS			ENTERTAINMENT		
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
TRANSITION MARKERS	262.63	172.90	435.53	272.52	201.12	473.64	202.33	217.62	419.95
FRAME MARKERS	48.97	98.13	147.10	25.95	94.22	120.17	72.06	58.39	130.45
ENDOPHORIC MARKERS	4.45	9.35	13.80	14.06	10.87	24.93	5.54	5.31	10.85
EVIDENTIALS	4.45	0.00	4.45	3.24	1.81	5.06	2.77	0.00	2.77
CODE GLOSSES	11.13	9.35	20.47	12.98	16.31	29.28	11.09	2.65	13.74

Table A.2: India: frequency of *interactional* markers (normalized)

	SPORTS			POLITICS			ENTERTAINMENT		
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
HEDGES	220.34	196.26	416.60	173.03	219.24	392.27	227.27	273.35	500.63
BOOSTERS	33.39	60.75	94.13	69.21	61.61	130.82	55.43	47.77	103.20
ATTITUDE MARKERS	213.67	327.10	540.77	107.06	159.45	266.51	135.81	175.16	310.97
SELF MENTIONS	82.35	84.11	166.46	38.93	123.21	162.14	11.09	53.08	64.17
ENGAGEMENT MARKERS	80.12	74.77	154.89	73.54	119.59	193.12	155.21	100.85	256.06

A.2 Quantitative Analysis of UK Comments

Table A.3: UK: frequency of *interactive* markers (normalized)

	SPORTS			POLITICS			ENTERTAINMENT		
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
TRANSITION	323.80	225.52	549.33	257.67	238.15	495.82	311.67	230.95	542.61
FRAME MARKERS	45.83	34.49	80.32	69.95	78.60	148.56	33.33	30.79	64.13
ENDOPHORIC MARKERS	10.96	15.92	26.88	12.82	9.36	22.18	31.67	11.55	43.21
EVIDENTIALS	4.98	0.00	4.98	2.33	6.08	8.41	0.00	3.85	3.85
CODE GLOSSES	23.91	29.19	53.10	27.98	25.27	53.25	13.33	23.09	36.43

Table A.4: UK: frequency of *interactional* markers (normalized)

	SPORTS			POLITICS			ENTERTAINMENT		
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
HEDGES	333.77	347.57	681.34	376.59	281.66	658.25	290.00	219.40	509.40
BOOSTERS	72.73	76.94	149.67	86.28	87.49	173.77	56.67	53.89	110.55
ATTITUDE MARKERS	198.27	108.78	307.05	146.90	128.67	275.57	146.67	192.46	339.12
SELF MENTIONS	1.99	18.57	20.57	3.50	103.40	106.90	1.67	11.55	13.21
ENGAGEMENT MARKERS	147.45	106.13	253.58	209.86	162.82	372.69	100.00	127.02	227.02

A.3 Quantitative Analysis of USA Comments

Table A.5: USA: frequency of *interactive* markers (normalized)

	SPORTS			POLITICS			ENTERTAINMENT		
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
TRANSITION	265.97	235.99	501.96	209.58	225.66	435.24	186.15	187.17	373.32
FRAME MARKERS	49.13	93.75	142.88	21.01	34.79	55.80	41.28	26.74	68.02
ENDOPHORIC MARKERS	8.47	12.93	21.40	9.66	9.81	19.47	4.67	0.00	4.67
EVIDENTIALS	3.39	7.54	10.93	4.54	3.12	7.67	4.67	2.06	6.73
CODE GLOSSES	25.41	49.57	74.98	18.17	42.37	60.54	24.92	26.74	51.66

Table A.6: USA: frequency of *interactional* markers (normalized)

	SPORTS			POLITICS			ENTERTAINMENT		
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
HEDGES	281.21	331.90	613.11	191.40	296.12	487.53	254.69	283.83	538.53
BOOSTERS	79.62	14.01	93.63	71.56	86.52	158.08	61.53	74.04	135.57
ATTITUDE MARKERS	135.52	144.40	279.92	132.33	170.81	303.14	115.27	121.35	236.62
SELF MENTIONS	37.27	65.73	103.00	48.84	60.21	109.05	48.29	43.19	91.48
ENGAGEMENT MARKERS	54.21	46.34	100.55	40.89	49.06	89.95	59.97	43.19	103.17

A.4 Domain wise analysis of comments(normalized)

Table A.7: Domain wise frequency of *interactive* markers with respect to political ideology

	SPORTS			POLITICS			ENTERTAINMENT		
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
TRANSITION MARKERS	293.64	224.50	518.15	237.65	228.33	465.97	222.30	207.52	429.82
FRAME MARKERS	47.47	79.66	127.14	34.15	60.43	94.58	44.10	38.30	82.40
ENDOPHORIC MARKERS	8.81	13.17	21.98	11.57	9.73	21.31	12.03	4.45	16.48
EVIDENTIALS	4.40	4.61	9.01	3.67	4.26	7.93	3.12	1.78	4.90
CODE GLOSSES	21.53	38.84	60.38	19.19	32.04	51.23	19.60	17.81	37.41

Table A.8: Domain wise frequency of *interactional* markers with respect to political ideology

	SPORTS			POLITICS			ENTERTAINMENT		
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
HEDGES	293.64	316.68	610.32	231.44	281.25	512.69	259.72	265.41	525.13
BOOSTERS	66.07	36.21	102.28	74.51	84.15	158.66	59.25	60.56	119.81
ATTITUDE MARKERS	183.53	161.30	344.83	129.27	151.27	280.54	126.97	155.86	282.83
SELF MENTIONS	29.85	56.62	86.47	35.28	85.98	121.26	29.85	39.19	69.04
ENGAGEMENT MARKERS	105.71	65.18	170.89	90.32	106.26	196.57	85.98	81.94	167.92

Table A.9: Domain wise frequency of *interactive* markers with respect to country

	SPORTS			POLITICS			ENTERTAINMENT		
	INDIA	UK	USA	INDIA	UK	USA	INDIA	UK	USA
TRANSITION MARKERS	435.53	549.33	501.96	473.64	495.82	435.24	419.95	542.61	373.32
FRAME MARKERS	147.10	80.32	142.88	120.17	148.56	55.80	130.45	64.13	68.02
ENDOPHORIC MARKERS	13.80	26.88	21.40	24.93	22.18	19.47	10.85	43.21	4.67
EVIDENTIALS	4.45	4.98	10.93	5.06	8.41	7.67	2.77	3.85	6.73
CODE GLOSSES	20.47	53.10	74.98	29.28	53.25	60.54	13.74	36.43	51.66

Table A.10: Domain wise frequency of *interactional* markers with respect to country

	SPORTS			POLITICS			ENTERTAINMENT		
	INDIA	UK	USA	INDIA	UK	USA	INDIA	UK	USA
HEDGES	416.60	681.34	613.11	392.27	658.25	487.53	500.63	509.40	538.53
BOOSTERS	94.13	149.67	93.63	130.82	173.77	158.08	103.20	110.55	135.57
ATTITUDE MARKERS	540.77	307.05	279.92	266.51	275.57	303.14	310.97	339.12	236.62
SELF MENTIONS	166.46	20.57	103.00	162.14	106.90	109.05	64.17	13.21	91.48
ENGAGEMENT MARKERS	154.89	253.58	100.55	193.12	372.69	89.95	256.06	227.02	103.17

A.5 Political Ideology Analysis of Comments (normalized)

Table A.11: Based on Political ideology: frequency of *interactive* markers (normalized)

	LW				RW			
	INDIA	USA	UK	Overall	INDIA	USA	UK	Overall
TRANSITION MARKERS	737.48	661.69	893.13	2292.31	591.64	648.82	694.62	1935.08
FRAME MARKERS	146.98	111.42	149.12	407.52	250.74	155.27	143.89	549.90
ENDOPHORIC MARKERS	24.05	22.80	55.45	102.303	25.53	22.74	36.82	85.09
EVIDENTIALS	10.47	12.61	7.31	30.39	1.81	12.72	9.93	24.47
CODE GLOSSES	35.19	68.51	65.23	168.93	28.31	118.67	77.55	224.53

Table A.12: Based on Political ideology: frequency of *interactional* markers (normalized)

	LW				RW			
	INDIA	USA	UK	Overall	INDIA	USA	UK	Overall
HEDGES	620.64	727.31	1000.35	2348.31	688.86	911.85	848.64	2449.35
BOOSTERS	158.03	212.71	215.67	586.42	170.12	174.06	218.32	563.02
ATTITUDE MARKERS	456.54	383.13	491.84	1331.51	661.71	436.55	429.90	1528.17
SELF MENTIONS	132.37	134.40	7.15	273.93	260.40	169.13	133.52	563.05
ENGAGEMENT MARKERS	308.87	155.08	457.32	921.27	295.20	138.59	395.97	829.76

A.6 Cross-culture Analysis of Comments (normalized)

Table A.13: Cross-cultural: domain wise frequency of *interactive markers* (normalized)

		TRANSITION FRAME		ENDOPHORIC	EVIDENTIALS	CODE
		MARKERS	MARKERS			
INDIA	SPORTS	435.53	147.10	13.80	4.45	20.47
	POLITICS	473.64	120.17	24.93	5.06	29.28
	ENTERT.	419.95	130.45	10.85	2.77	13.74
	Overall	1329.12	397.72	49.58	12.28	63.50
USA	SPORTS	501.96	142.88	21.40	10.93	74.98
	POLITICS	435.24	55.80	19.47	7.67	60.54
	ENTERT.	373.32	68.02	4.67	6.73	51.66
	Overall	1310.51	266.70	45.54	25.33	187.18
UK	SPORTS	549.33	80.32	26.88	4.98	53.10
	POLITICS	495.82	148.56	22.18	8.41	53.25
	ENTERT.	542.61	64.13	43.21	3.85	36.43
	Overall	1587.76	293.01	92.28	17.24	142.77

Table A.14: Cross-cultural: domain wise frequency of *interactional markers* (normalized)

		HEDGES	BOOSTERS	ATTITUDE	SELF MEN-	ENGAGEMENT
				MARKERS	TIONS	MARKERS
INDIA	SPORTS	416.60	94.13	540.77	166.46	154.89
	POLITICS	392.27	130.82	266.51	162.14	193.12
	ENTERT.	500.63	103.20	310.97	64.17	256.06
	Overall	1309.50	328.15	1118.25	392.77	604.08
USA	SPORTS	613.11	93.63	279.92	103.00	100.55
	POLITICS	487.53	158.08	303.14	109.05	89.95
	ENTERT.	538.53	135.57	236.62	91.48	103.17
	Overall	1639.16	387.28	819.68	303.53	293.66
UK	SPORTS	681.34	149.67	307.05	20.57	253.58
	POLITICS	658.25	173.77	275.57	106.90	372.69
	ENTERT.	509.40	110.55	339.12	13.21	227.02
	Overall	1848.99	434.00	921.74	140.68	853.29

Table A.15: Cross-cultural: political ideology wise frequency of *interactive markers* (normalized)

	INDIA			USA			UK		
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
TRANSITION	737.48	591.64	1329.12	661.69	648.82	1310.51	893.13	694.62	1587.76
FRAME MARKERS	146.98	250.74	397.72	111.42	155.27	266.70	149.12	143.89	293.01
ENDOPHORIC MARKERS	24.05	25.53	49.58	22.80	22.74	45.54	55.45	36.82	92.28
EVIDENTIALS	10.47	1.81	12.28	12.61	12.72	25.33	7.31	9.93	17.24
CODE GLOSSES	35.19	28.31	63.50	68.51	118.67	187.18	65.23	77.55	142.77

Table A.16: Cross-cultural: political ideology wise frequency of *interactional markers* (normalized)

	INDIA			USA			UK		
	LW	RW	Overall	LW	RW	Overall	LW	RW	Overall
HEDGES	620.64	688.86	1309.50	727.31	911.85	1639.16	1000.35	848.64	1848.99
BOOSTERS	158.03	170.12	328.15	212.71	174.57	387.28	215.67	218.32	434.00
ATTITUDE MARKERS	456.54	661.71	1118.25	383.13	436.55	819.68	491.84	429.90	921.74
SELF MENTIONS	132.37	260.40	392.77	134.40	169.13	303.53	7.16	133.52	140.68
ENGAGEMENT MARKERS	308.87	295.20	604.08	155.08	138.59	293.66	457.32	395.97	853.29