Università degli Studi di Roma "La Sapienza"

Department of European, American and Intercultural Studies

Yearning for Identity in the Glimmer of America: A Corpus-assisted Analysis of the South Asian American Diaspora Literature

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Department of the European, American and inter-cultural

studies

by:

Soheila Heidari

Supervisor:

Professor Giorgio Mariani

Advisor:

Dr. Federica Perazzini

Yearning for Identity in the Glimmer of America: A Corpus-assisted Analysis of the South Asian American Diaspora Literature

By:

Soheila Heidari

Acknowledgements

This dissertation would not have been possible without the guidance of some wonderful individuals whose contributions have been pillars of support for the completion of this study. It is a pleasure to convey my gratitude to them all in my humble acknowledgment.

First and foremost, I would like to offer my sincere thanks to my dissertation supervisor, Professor Georgio Mariani, to whom I owe an enormous debt of gratitude. His constant support, constructive comments and suggestions throughout my research along with his invaluable helps in facilitating the administrative works as the Director of our Ph.D. program, are what I am forever grateful for. My special appreciation and thanks go to my amiable advisor, Dr. Federica Perazzini, for her constant flow of support and guidance. She was a source of encouragement for me, and her priceless suggestions improved the quality of this research.

I am deeply grateful to Professor Jaffer Sheyholislami at the School of Linguistics and Language Studies who kindly accepted to be my mentor during my visit to Carleton university. He has been beyond generous in sharing his immense knowledge and insightful suggestions that reshaped my thesis.

At this juncture, I should also express my appreciation to all the people who contributed to my personal and professional time at Carleton, especially the friendly staffs of the MacOdrum Library who were extremely helpful in the corpus collection process of my research.

Finally, I would like to thank my husband who has supported me in all my pursuits and has been a constant source of encouragement.

Abstract

The analysis of diaspora and its characteristics have constituted a substantial body of recent studies and researchers in different disciplines have tried to provide a new definition of diaspora. There are several reasons for the growing interest in the notion of diaspora, the most important of which is the possibility it provides to question the basic, yet fundamental concepts. Diaspora breaks the conceptual relations between concepts and their meanings and acts as a third space for them to be redefined all over again. Diaspora deconstructs the meaning of home, identity, culture, language, memories, ethnicity, belonging and traditions and by doing so, forces those in diaspora to re-evaluate the meanings, re-establish the connections and reread themselves using the newly defined concepts.

Given that diaspora provides a space for concepts to be in the process of becoming and not being, it also makes it possible for the stereotypes to be revisited and redefined. South Asian Americans are one of the most significant minority groups in the United States who not only contributed to the advancement of the Diaspora literature but also tried to redefine themselves and the existing stereotypes about them. This study aimed to uncover the central framework of the South Asian American Diasporic literature and to investigate the linguistic and literary patterns of it.

The evaluation was carried out through the employment of the corpus analytical methods, including keywords, concordance, and collocations on the compiled corpus for this study. The analyses of the corpus revealed the distinguishing features of the South Asian American Diaspora literature written by the female authors and identified the embedded ideas within the context of this group. This study also illustrated the similarities and the differences between the

findings of this study using corpus analytical methods and the literary readings. The findings of

this study can be helpful in identifying the benefits of computational criticism in the study of

literature.

Table of Contents

List of Figures

List of Charts

List of Tables

Introduction

The analysis of diaspora and its characteristics have constituted a substantial body of the recent studies and researchers in different disciplines have tried to provide a comprehensive picture of diaspora. The concept of diaspora which was basically used to refer to the experience of the Jews living in the exile, evolved over the time and it has gradually extended to the political, economic, cultural and social arenas. There are several reasons for the growing interest in the notion of diaspora, the most important of which is the possibility it provides to question the basic yet fundamental concepts. By breaking the established relation between the concepts and their meanings, diaspora acts as a space for redefining them. Diaspora poses challenges to the already defined concepts such as home, identity, culture, language, memories, ethnicity, belonging and traditions and by doing so forces those in diaspora to re-evaluate the meaning of them and re-establish the connection between the concepts.

Given that meanings are not taken for granted within diaspora, therefore, it can be claimed that diaspora studies allows to address and confront certain classified assumptions and stereotypes in that respect. South Asians are one of the significant ethnic minority group in the United States who not only contributed to the development of the Diasporic literature, but also attempted to challenge the existing stereotypes about themselves. By employing corpus analytical tools, this research aims to examine the literary and linguistic characteristics of the South Asian American Diaspora literature and to inspect whether they tend to challenge the existing concepts and provide new definitions. The employment of Corpus linguistics in literary texts helps to highlight the main characteristics of the text along with offering new insights into the corpus by demonstrating its overall structure. It also makes it possible to ask questions that could not be addressed using previous methodologies. This study sets out to investigate the

linguistic and literary characteristics of the South Asian American Diasporic literature. The answers to the following questions constitute the main part of this study:

 1. Does the employment of the corpus analytical tools uncover the central framework of South Asian American Diasporic literature?

 2. Is it possible to identify the linguistic and literary structure of the Diasporic literature?

 3. How have the authors adopted a certain terminology to convey their messages and whether the identification of these patterns determine new findings about diaspora ?

 4. Whether the findings of the corpus-driven approach add to the current literary views on the Diasporic literature or dismiss them?

 To provide a better understanding of the capabilities of digital humanities, the first chapter will provide an introduction to the history, development and potentials of digital humanities and its contributions to the digital literary studies. The corpus linguistics' essential functions in literary analysis will be examined as well. The second chapter will cover postcolonial literature and its connection with diasporic studies, along with offering the main theories and perspectives that each cover. South Asian American Diaspora will be explored afterward and the main characteristics of the South Asian American diaspora will be highlighted. Third chapter will also include the approaches of the female authors regarding the existing stereotypes about South Asian Americans and the way they react to these classified assumption in their works. After identifying Diaspora's main features and methods for decoding the meaning of diaspora in digital humanities, the fourth chapter will include the design, construction and the evaluation of the specific corpus of this research. The evaluation process will be explained in detail and the results will be provided. The quantitative results obtained in the previous section will be explored in the final chapter and the findings will be presented. The

results will be evaluated in respect to research questions and it will be determined if the results offer a new findings about the Diasporic literature and whether it is possible to propose a new perspective to diasporic literature based on the findings. This chapter will be concluded with future directions of the study and the suggestion for further studies.

Chapter One

Digital Humanities

Seen as a mere tool without any inherent analytical power of its own, the computer in literary studies enhances the critic's powers of memory electronically, thereby providing a complete database of findings that meet all predefined patterns or search criteria.
(Siemens and Schreibman, *A Companion to Digital Humanities*)

1.1 Introduction

This chapter covers the concept, usage, and perspectives toward digital humanities and digital literary studies. After providing a brief introduction on its emergence, the main reasons underlying the significance of digital humanities in contemporary era will be presented. The second section will include three stages of digital humanities, in terms of temporal and conceptual view. Accordingly, this section will be divided into past, present and future perspectives which are running in the main arguments of each epoch. In each part of the second section, there will be an attempt to answer the significant questions of that period, starting from how the move from humanities computing to digital humanities occurred in the past, to how the

digital tools, methodologies, and process could be employed now and where digital humanities

will go from here. After examining the emergence, developments and the prospects of digital

humanities that encompasses different disciplines, the following section will include digital

literature and the views and interpretations of literary critics in that respect. The necessity of

literature on a larger scale will be presented and the main arguments of critics in respect of

employing distant reading will be addressed. In the following section the significant studies that

have been carried out on digital literature so far will be presented and in the final section, the

conclusion of the discussions will be presented.

1.2 Digital Humanities; A Preliminary Overview

The advancements in the production and circulation of digital contents have given rise to

the proliferation of digital content, its dissemination, and expansion of interdisciplinary research

in digital humanities, and it has urged critics to provide a more precise definition of what

constitutes digital humanities and whether it should be in a singular or plural format.

There are several definitions for digital humanities, most of which include the use of

digital tools for research purposes. Elijah Meek defined digital humanities as:

> … the integration of sophisticated, empirical techniques utilizing tools and
>
> techniques typically associated with practical science into the study of
>
> traditional humanities questions. It represents a more exploratory and less
>
> quantitative approach than social sciences in the use of such tools, but it also
>
> represents ambitious attempts to model nuanced human wisdom in ways that,
>
> like early flying machines, are beautiful, quite impractical and often fail(qtd.

in *The Big Humanities: Digital Humanities/Digital Laboratories*, 16).

By addressing digital humanities as "scholarly primitive", Schreibman et al. refer to the common ground of this conceptual framework across disciplines as; discovering, annotating, comparing, referring, sampling, illustrating, and representing (*A New Companion to Digital Humanities, 151*). The book entitled *Digital_Humanities* defines it as "new modes of scholarship and institutional units for collaborative, transdisciplinary, and computationally engaged research, teaching, and publication." According to this definition, digital humanities "explores the universe in which print is no longer the primary medium in which knowledge is produced and disseminated" (*Digital_Humanities* ,122). The Day in the Life of the Digital Humanities, which is also addressed as the Day of DH, is a community publication project that gives voice to the distinguished digital humanists to define this evolving research area. Kathie Gossett, a researcher from Lowa University, addressed digital humanities as "interdisciplinary; by necessity, it breaks down boundaries between disciplines at the local (e.g., English and history) and global (e.g., humanities and computer sciences) levels." Mark Tebeau from Cleveland State University viewed digital humanities as a "collaborative, open, and emerging field of inquiry. A state of mind, a methodology, and theoretical approach to knowledge… like jazz, in that, it is about process, as well as the outcome" (*Debates in the Digital Humanities*, 68).

Potter defines digital humanities as the attempt to put two different arenas of technology and criticism into contact to form a new type of literary study that is consistant with scientific methods but still carries the values of the humanities (qtd.in *A Companion to DigitalHumanities*, 419). Anne Mangenin in *Reading on a Technological Platform* argues that both the reading process and digital text are influenced by the fact that the interaction between individuals and texts happen within a distance from the actual text, while in reading a physical book, a close

interaction is always present (405). However, many scholars do not consider the absence of close interaction between the physical text and the reader as a drawback for digital humanities and instead, they believe this distance could foster a vibrant dialogue between the digital world and the real world. Accordingly, they believe that if an author is writing a fiction, beside the actual act of writing, he or she is investigating different types of textual navigations and interplays. As specified by Scott Rettberg in *Electronic Literature as Digital Humanities*:

> The author of a short story producing a narrative generator is not only writing a story, but also collaborating with the research and development wing for new literary forms that engage with technology on an aesthetic level. New works not only function as individual artistic expressions but also as documented experiments in applied technology and formal innovation. Writers working in new media are both creating discrete literary experiences and testing the chemistry of a particular creative admixture of writing and technological apparatus (129).

However, the number of digital humanists who refuse to consider digital humanities as a new research field in its traditional sense is constantly increasing and instead, they acknowldge digital humanities as an umbrella term for a group of experimental approaches that came together to address the fundamental questions of humanities including the meaning of beauty or truth or how the given social circumstances get translated into aesthetic objects. They believe that digital humanities is an attempt to collect heterogenous phenomena under a unified framework to examine the underlying meaning of them and evaluate the transformation they have gone through. By bearing the conceptual scope of the digital humanities, David Goldberg has divided digital approaches into four groups; the first group is concerned with the advancement of the digital tools ranging from visualization and data mining tools to search engines and publishing

platforms. The second group includes humanistic projects that can be realized through the support of digital tools. They may not need digital tools, but a better result will be achieved if digital tools are employed. The third category contains critical thinking about digital context which is considered as the object of the analysis as well. The last category includes a combination of practices that are identified as digital humanities including projects that are simultaneously humanities-driven and digital-driven ("Deprovincializing Digital Humanities", 163-164).

Over the past two decades, humanities have been subjected to radical transformations embedded in the significant progressions in the core humanities. The emergence of digital humanities can be traced back to the 1940s, when the compelling impact of computing was observed on different research branches. According to Stephen Ramsay, although digital humanities has undergone a long process of revolutionary transformation to be in line with the society, it has a precise indication of its founder and the process through which it is formed (*Reading Machines: Toward An Algorithmic Criticism,*1).

In the late 1940s, an Italian Jesuit priest, Father Roberto Busa carried out the first attempt to automatically generate the concordance of the works of Thomas Aquinas and related authors through a computer. The profound impact of his work not only gave rise to humanities computing, but also transformed the perspective toward the humanities altogether. By the 1960s, other researchers had started to appreciate the advantages of analyzing concordance in their research and a series of individual and institutional articles were published on this subject, the most important of which can be referred to the study of the Federalist Papers by Mosteller and Wallace (qtd. in *A Companion to Digital Humanities,* 22).

So, at the mid-stages, there had been a continuous emphasis on the linguistic analysis within the humanities-computing, but in the 1980s and 90s the emphasis moved toward digitized editions of the books and their online availability, including the works of individual authors as well as periodic or thematic archives. The projects, including Walt Whitman Archive 1995, Perseus Digital Library Project 1985, and the Rossetti Archive 1993 were among the most important projects that started during that period (Iris Gemeinbock, 30). This process has continued to present and as Svensson stated, in the last 20 years, there has been a significant progress in funding digital humanities' projects, initiatives and centers worldwide (Svensson, 39).

1.3 Digital Humanities; Temporal and Conceptual Prospects

The continuing transformation of digital humanities, either conceptually or functionally, can be broken down into three phases that respectively cover its emergence in the past, its current progress and its potential evolvements in the future. Since its rise, digital humanists have worked hard to uncover the main issues or concerns, either by providing answers to the issues that have been raised so far in humanities or raising new ones that have not been asked before and providing the platform for further discussions or debates in the future. Three temporal and conceptual phases of digital humanities include; early, current and future prospects that are addressed in the following.

1.3.1 Early Prospects

Digital humanities, which was addressed as humanities computing upon its emergence, was mostly concerned with employing tools for linguistic purposes. Yet, a skeptical atmosphere

surrounded the exploitation of tools for the study of humanities at the early stages. Kim Jensen argues that at the rise of digital humanities, when the analytical techniques were rather primitive, "doing computer-aided linguistic analysis was genuinely frowned upon by the academic establishments and it was viewed merely as the work of a spanner-wielding handyman as opposed to the proper academic work of the rationalist elite" ("Linguistics and the digital humanities", 57). Rosanne Potter, the author of "Literary Criticism and Literary Computing" also connotes the same hesitating perspective toward digital humanities saying that "until everything has been encoded, or until encoding is a trivial part of the work, the everyday critic probably will not consider computer treatments of texts" (93).

Other discussions of the early period of digital humanities involved the discussions on the move from humanities computing to digital humanities, how the process happened and what kind of effects it had on the research arena. According to Anne Burdick, humanities computing emerged at the intersection of computing or digital technologies and humanities, and in short, it included new methods for conducting scientific researches which had two main features that are still authentic in digital humanities; first, digital humanities was considered as the continuation of the same traditional knowledge but the fundamental source of its production and contribution was not words but codes. Second, there was a two-way relationship between the humanities and the digital world; that at the same time that the scientific branch employed the technology for research, it criticized its impact on the cultural heritage and digital culture as well (*Digital_Humanities,* 49-52).

1.3.2 Current Prospects

Current debates on digital humanities are mostly concerned with the adaptation of computational tools for scholarly examinations as well as discussion on moving from the traditional approach of focusing on the physical texts and books into the digital approach in which the print is no longer the core. For example, by focusing on the historical trend of digital humanities up to present, Nancy Ide sought to authorize digital humanities and she claimed that in 1960s, a text-based discussion began to emerge on a marginal level, but was gradually extended to other areas of culture, informatics and computer science, and eventually the discussion came to a head in the 1990s by the advent of the World Wide Web and the rise of personal computing as the defining public space and in that context, digital humanities became a transformational area of experimental scholarship within the humanities disciplines ("Preparation and Analysis of Linguistic Corpora", 287*).*

Although digital humanities have been around for more than four decades, there are still discussions about what it really means or how it should be defined. Yet, some digital humanists argue that digital humanities' definition has been provided in excess; therefore, further explanations should be avoided. But as much as they push the question back, it repeatedly emerges in the larger discussions. The importance of redefining knowledge states an important viewpoint that a lot of digital humanists share. At the core of these discussions is the idea that humanities scholarship can be advanced in a digital environment and it does not need to be limited by the printed codex, therefore, administrative structures are required to pass through the print process and article formats (*A New Companion to Digital Humanities,* 462). And consequently, digital scholars have come to this agreement that there is a profound necessity to redefine and reinterpret knowledge to include the newly formed methodologies, knowledge, and

fields. The biggest argument in this respect is whether to include Zotero as a scholarly work or as a scholarly service. William G. suggests that there might be a time that humanities scholars will be able to bridge the gap between the digital approach and the core humanities to not only include their digital-based research in the ongoing discussions in humanities but also to form the digital scholarship ("The Promise of the Digital Humanities and the Contested Nature of Digital Scholarship"*, 525*). Hughes in *Digital Collections: Use, Value and Impact* argues that employing digital contents, tools and methods have given rise to studies that have the capability of changing the infrastructure of understanding in general and they can facilitate the emergence of new knowledge through one of the following ways;

• Firstly, by facilitating and enhancing the existing research, by making research processes easier via the use of computational tools and methods.

• Secondly, by enabling research that would be impossible to undertake without digital resources and asking new research questions that are driven by insights only achievable using new tools and methods(123-5).

An important part of the current discussion in digital humanities revolves around the necessity of introducing large-scale scholarly centers, journals, curriculums and projects with the smaller ones, and to also provide the less known centers, scholars and projects with a chance to be included in the main discussions of the field. In this perspective, the ongoing collaboration between the representatives of different centers, researchers and critics of different genres plays a significant role, beside including newly founded centers around the world beside that of the United States, Europe, and Canada. According to Schreibman, Consortium for Electronic Literature can be considered as one of the best-known digital humanities establishments for research collaborations in the electronic literature. Consortium for Electronic Literature was

formed to bring the international actors of the field, particularly the organizations and projects that develop research databases and archives, into closer contact with one another to support communication and to establish machine-level interactions between databases and archives (*A New Companion to Digital Humanities, 135).*

Closely connected to the former issue is the investment in the formation and collection of digital contents and international initiatives that in return, support the creation, management and the protection of digitalized content in a larger scope. The development of this process has given the chance to researchers to pose questions that could not be asked earlier but made possible because of the existence of digital tools and methodologies. For instance, according to Houghton, now, it is possible to investigate the exact way through which a researcher has used digital contents and evaluate his methodological practice (Houghton, 147). Schreibman et.al. in *The Digital Humanities and Humanities Computing: An Introduction,* share the same perspective toward establishing digital formats along with physical versions of them and argue that:

> Although the breadth of the covered fields is wide, what is revealed is how computing has cut across disciplines to provide not only the tools, but also the methodological focal points. There is, for example, a shared focus on preserving physical artifacts (written, painted, carved, or otherwise created), which is left to us by chance (ruin, and other debris of human activity), or that which has been near-impossible to capture in its intended form (music, performance, and event). Yet, many disciplines have gone beyond simply wishing to preserve these artifacts, what we might now call early forms of data management, to re-represent and manipulate them to reveal properties and traits not evident when the artifact was in its native form. Moreover, digital humanities now also concern itself with

the creation of new artifacts which are born digital and require rigorous study and understanding in their own right (2).

1.3.3 Future Prospects

Other scholars who agree with Matthew K. Gold that the current time can be called as "digital humanities moment"(*Debates in the Digital humanities*, IX), direct their attentions toward the necessary steps that are required to take the digital humanities to the next stage and in doing so, what kind of opportunities will be available for scholars. Respectively, McGann predicted that "in the next 50 years, the entirety of our inherited archive of cultural works will have to be re-edited within a network of digital storage, access, and dissemination" (2).

Gold, in his introduction to *Debates in the Digital Humanities,* pinpointed the fact that major publications including the Boston Globe, New York Times, Nature, and the Chronicle of Higher Education have highlighted digital humanities not just "the next big thing," as the Chronicle of Higher Education claimed in 2009, but simply "the Thing," as the same publication noted in 2011" (*Debates in the Digital Humanities* , ix).

Consequently, more researchers have concentrated on the affinity between human science and new tools that can cultivate new perspectives. Nancy Idle argues that with the advent of digital technology, data storage of the countless books and their evaluations have been facilitated and new analytical tools have made it possible to address the questions that were posed by modern communication and in this respect, new methodological tools should be designed and used to evaluate the connection of digital era and material culture ("Preparation and Analysis of Linguistic Corpora", 289). This connection could be made when additional spaces are created within the text to include further related discussions. Marcel Pope addressed the importance of forming intervals within the advanced multilayered structure of the text, so that outside contents

could be attracted into the text and by doing so ideological thinking could be animated (*New Literary Hybrids in the Age of Mltimidea Expresion*, 1).

Along with other persectives toward the future of digital humanities and the directions it should follow, some humanities scholars have a utopian perspective toward digital humanities, and on that basis, they believe that the idea of digital humanities helps to move across traditional boundaries of disciplines. For instance, Mark Marino from University of Southern California, addressed digital humanities as "a name that marks a moment of transition, a temporary epithet for what will eventually be called merely humanities" (qtd. in *, Debates in the Digital Humanities,* 68*).*

The final prospect regarding the future of digital humanities presented here is to employ algorithmic evaluation as a new science that can go beyond the defined limitations. Stephen Ramsay has proposed employing a method called "algorithmic criticism" for analyzing humanities and he addressed it as a hermeneutical foundation, but he argued that computer is incapable of offering "the shift to a redemptive worldview"; therefore, the humanists should take this philosophical step to employ algorithmic criticism (*Reading Machine Toward an Algorithmic Criticism,* 9). He argued that, prior to algorithmic criticism, it was the computational text analysis that was considered as the most scientific form of the literary investigation, which incorporated advanced technical tools into the analysis of a text, but algorithmic criticism is derived from algorithmic manipulation of the text to penetrate into the core of literary studies (2).

1.4 Panoramic Overview of Digital Literary Studies

The preliminary studies of literature through the computer in the 1960s and 70s included the identification of the patterns and strings within the electronic texts, tied with word list and

concordance of the texts that were initially available in the format of a book and then digitalized. At this stage, the precise details of the significant literary texts were made available online and as John Burrows highlighted, researchers refused to bound themselves to the traditional rules of the literary criticism. He addressed this issue by saying that "It is a truth not generally acknowledged that, in most discussions of works of English fiction, we proceed as if a third, two-fifths, a half of our material were not really there" (*A Computation into Criticism. A Study of Jane Austen's Novels and an Experiment in Method,* 1).

This transitional period brought about dramatic changes that resulted in the emergence of digital humanities and Michel Foucault's *'dream of a new age of curiosity'* was realized. David Dowings andSosnoski addressed this era as "a moment when the modes and the technology for cultural reproduction are shifting, this time from print to electronic environments which opens new possibilities for freedom as well as oppression" ("As the CultureTurn: Postmoden *Works and Days",10*). Susan Hockey in *Electronic Texts in the Humanities: Principles and Practice*, explains the role of digital tools in the literary studies as quoted below:

> Computers can assist the study of literature in a variety of ways, some more successful than others ... Computer-based tools are especially good for comparative work, and here some simple statistical tools can help to reinforce the interpretation of the material. These studies are particularly suitable for testing hypotheses or for verifying intuition. They can provide concrete evidence to support or refuse hypotheses or interpretations which have been based on human reading and the somewhat serendipitous noting of interesting features (66).

Electronic literature, that has come to be used by the broader field of digital scholarship, resists any fixed definition, yet some features of it can be distinguished. The Electronic

Literature Organization included hypertext fiction, poetry, and kinetic poetry as the types of electronic literature and defined it as "works with important literary aspects that take advantage of the capabilities and contexts provided by a stand-alone or networked computer" (*Electronic Literature).* Katherine Hayles illustrates electronic literature as "a first-generation digital object created on a computer and (usually) meant to be read on a computer" (*Electronic Literature, 3*). However, Scott Rettberg provides a more comprehensive definition of electronic literature by referring to it as an umbrella term to connote different forms of literary practice that benefit from computational and collaborative features of computers for the digital evaluation of the narrative or poetic features of the text hypertext fiction and kinetic multimedia ("Electronic Literature as Digital Humanities*",* 127). Retberg defines electronic literature as a subcategory of digital humanities that works on different levels and includes:

   1. creative digital media practice in electronic including

literature;

   2. the development of specific platforms for creative practices in digital media;

   3. theoretical work and analysis works of electronic literature to build new understandings of contemporary textuality and "digital vernaculars";

   4. The establishment of the networked scholarly practices, digital publications, research infrastructures, and social networks particular to the digital media research environment;

   5. meta-analysis and visualization research based on electronic literature metadata (128).

    In general, three levels of scholarly practices are recognized in both the broader realm of digital humanities as well as the electronic literature distinguished by theory and analysis, toolmaking, platform development, and applied research ("Electronic literature as Digital Humanirtes", 127). Unlike other humanities that can be characterized by their generic, temporal

and regional features, the boundaries of emergent fields of digital humanities tend to be more fluid. Rettberg argues that the point of distinction between electronic literature and digital humanities is that the earlier explores the impacts of the computational tools and networks solely on literary texts, but the later one is broader in the sense that its focal point is on the digital methods for the research within literature, history and other established humanities disciplines. However, the common point between them is that neither of them is defined with its connection to a specific genre or a historical era, but instead, they are described with their engagements with digital tools and technologies ("Electronic Literature as Digital Humanities", 127).

Also, given that the digital study of literature provides innovative insights into the style and literary theory that would not be accessible without employing tools; these studies have managed to locate themselves in the context of meta-discourse and the question of the method remains at the core of the digital study of literature (Thomas Rommel, 87).

1.4.1 Literature at scale

The adaptation of the computational approach has been greatly beneficial in evaluating literary texts in a large scale which could not be studied otherwise. The expansion of the scope of the texts that made available online in digital format, along with the development of the specialized tools for the literary examinations, not only transformed the methodology of research and the type of questions that could be asked and examined but also enlarged the scope of the context that could be evaluated. For instance, researchers are provided with the chance to investigate the prospects of automatic computational genre-attribution, to examine the authorship of a writer, compare his works with others or to evaluate the whole genera or period (Burrows 1987; Stubbs 2005). This change in the methodology and perspective was documented by the

Stanford literature scholar, Franco Moretti in 2000 when he offered the term "distant reading" in his article in the New Left Review titled as "Conjectures on World Literature." Moretti suggested distant reading as an irony to "close reading," which is the standard method of literary study, devoted to a detailed examination of specific chapters, passages, and sentences in single texts. He argued that since the focal point of the close reading is on a few, typically canonical texts, close reading is not capable of responding to the necessity of evaluating the whole national literature and he alternatively proposed distant reading that could enable scholars to "focus on units that are much smaller or much larger than the text: devices, themes, tropes—or genres and systems"(15). Moretti suggests that employing computer for the study of hundreds of texts at once gives the access to deeper meanings and insights that would not have been physically possible using traditional methods. However, it is important to note that distant for Moretti, as he puts it is "*a condition of knowledge*" (57).

The rise of distant reading can be considered as one of the significant impacts of digital humanities on literature. As stated by Rauscher, the promising possibilities of computational approaches to literature, language and the humanities and the emergence of distant reading as the new methodology for literary studies with its quantifiable aspects, provides the context for further research on new methods of analysis in social and spatial aspects of the texts and helps to iswntify the structures and patterns across a large number of novels ("Grasping cities through literary representations", 70). Amir Khadem, in his essay entitled "Annexing the unread: a close reading of "distant reading ", describes distant reading as:

> As a method of enquiry based on tracing a formal element through a vast body of
> works in a historical and geographical span, and then trying to build an
> explanatory model of the emergence, demise, or transformation of certain aspects

of literature. It tends to focus on both canonical and non-canonical literature and aims to find new correlations in literary history by embracing works of literature that are usually neglected. To describe these neglected works, Moretti uses the term ''the great unread,'' borrowed from Margaret Cohen (1).

Amir Khadem further divides distant reading into two distinct methods of ''anatomical distant reading,'' and "epistemological distant reading'' where in the earlier, a text plays a minor role and the researcher focuses on large-scale patterns of change over the time in the literary field and in the later, the researcher avoids the analytical study of the work as a whole, but rather focuses on a precise perspective within a huge number of texts, to trace a limited specified elements in them. (Khadem, 415) From both viewpoints, however, Moretti considers distant reading as not "an obstacle, but a specific form of knowledge" (Moretti, 1).

It is important to note that distant reading goes hand in hand with data mining and visualization. Text mining is a far-reaching quantitative approach with an interdisciplinary nature that is mostly employed in projects in computer science, statistics, linguistics, sociology, and other social sciences. In return, visualization makes the hidden patterns across the data visible and traceable that would have otherwise remained unexplored. Visualization also equips the researcher with the roadmap to perceive connections and relationships between those abstracted elements and progress toward new interpretations and understandings.

In this regard, the best examination of a text, according to Janneke Rauscher, includes managing different degrees of distance and moving between close and distant readings refered as an iterative process ("Grasping Cities through Literary Representations. A Mix of Qualitative and Quantitative Approaches to Analyze Crime Novels," 79). This process gathers the quantitative results of the first phase and quantifies and refines the findings by mining the data.

Then, the data cpuld be approached from a mid-level distance by employing the qualitative

content analytical tools. The following figure is the illustration of the defined process by

Rauscher:

Figure 1.1 Overview of Multi-Phase Research Design, "Grasping cities through literary
representations: a mix of qualitative and quantitative approaches to analyze crime novels"

Phase 1   **Phase 2**

| **Gathering and**
| **preprocessing** |

| **Close reading**
Sample novels
dimensions of comparison
and first hypotheses |

| **Computer-aided QCA**
Mid-level distance
and phenomena |

| **Close reading**
Of automatically retrieved
sentences and passages |

| **Distant reading**
Quantitative analysis of micro-
level lexical and semantical
patterns |
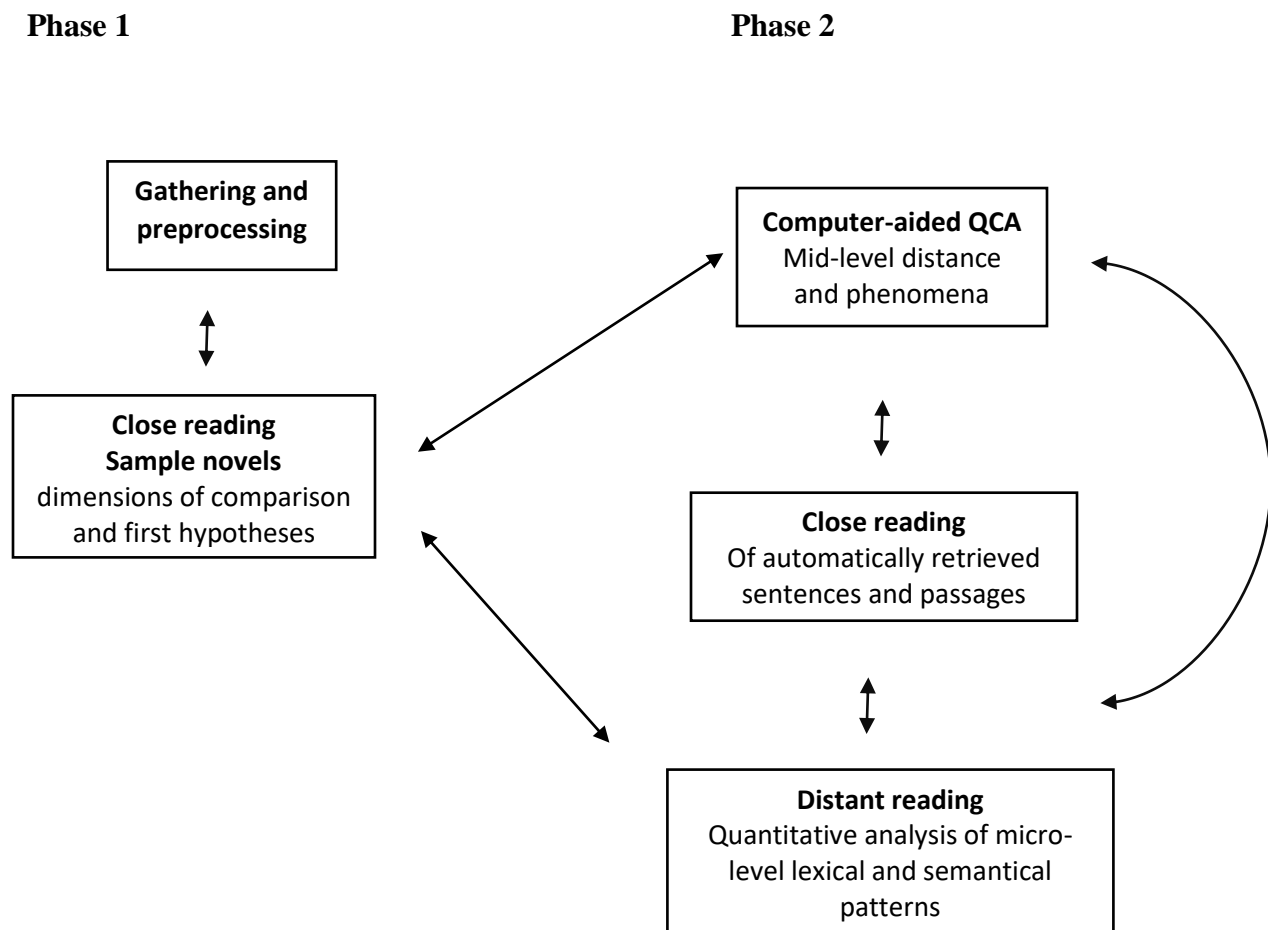
1.5 Digital humanities and Significant Literary Studies: Example of literary works

The most distinguished studies that have been carried out through quantitative and

qualitative methods so far in respect to literary texts are concerned with corpus stylistic,

discourse analysis, authorship as well as identifying keywords and their concordance and collocations to address, discuss or reject the main arguments of the text. In the following, some of these significant studies and the issues they have discussed are presented. However, it should be mentioned that the listed categories are not necessarily separate from one another and based on the research questions, or the scope of the research, either of them or all of them could be employed together. The following distinction is made solely for further clarification of the compelling research that have been carried out therein.

1.5.1 Corpus linguistic and corpus stylistic

Corpus is a fundamental part of any research on language and its features. The rise of computing tools has facilitated data collection in electronic format that could be examined for different purposes, including language or literary features, certain characteristics or descriptive properties. The golden era of the linguistic corpora started in the 1990s, which continued to the present day. In this regard, Sara Kerr has examined the features of a large corpus on the novels written, between 1640 and 1830 and compared the resulted with the novels of Jane Austin and Maria Edgeworth.

However, many literary scholars have acknowledged the deficiency of the specialized literary corpora that could address their concerns and they agreed that many of the available corpora were compiled only for linguistic investigation. Therefore, a large part of the current research either highlights this concern or sets the tone for the necessity of compiling the required corpora individually or collectively. C18P is a 9.7 million-word corpus of eighteen century prose fiction designed specifically for literary purposes. The corpus included short fictions and early development of the novels up to the Victorian era and it has been described to be helpful in

analyzing specific linguistic features of the included literary genres and forms.

One of the significant arguments underlined by many of the critics regarding Conrad was the subject of ambiguity and vagueness in both his style of writing as well as his narration in *Heart of Darkness.* Stubbs Made a comparison between the frequently used phrases in *Heart of Darkness* and those that occurred in the 100-million word British National Corpus and ulitmately confirmed the argument about vagueness in this novel by claiming that the expressions in *Heart of Darkness* are "abstract and extremely vague and acquire evaluative connotations" (Svenja Adolphs, 65).

As mentioned, corpus studies can avail themselves for comparing determined authors, texts or features with a bigger corpus to identify the distinguishing characteristics of each. This can be done by generating key words by comparing the frequencies of the words in the text with the frequencies of those words in a reference corpus. For instance, Jonathan Culpeper set the corpus of *Romeo and Juliet* against all of the plays by Shakespeare to examine the use of words through the analysis of individual characeters' speeches. Other examples include the study of particular authors such as Dickens or Austin or the style of an author within a special text such as *Eveline* or *Heart of Darkness*.

In literary studies, corpus can also be used to examine a particular idea or hypothesis. For instance, Paul Baker, in his book entitled *Using Corpora to Analyze Gender,* benefited from corpus analysis to investigate the representation and differences of gender and sexuality between men and women in the English language. He intended to convey this message that without requiring to be a computer or math wizard, critics can achieve a lot by employing Corpus Linguistics (6).

1.5.2 Concordance, collocation, and frequency

For information on the repeated words or structures reveals a lot about the text, analyzing the frequency of a word, expression or idiom is at the core of corpus linguistics. In this respect, Burrows, in *Computation into Criticism: A Study of Jane Austen's Novels,* explored the frequency profile of the individual terms in relation to idiolects of special characters in Jane Austin's novels. Also, Giuseppina Balossi, in an attempt to identify literary language and characterization style of Virginia Woolf, examined *The Waves*, which is considered as the most experimental work of Virgina Wolf. The novel includes the stories of six friends from childhood to adulthood in a series of monologue, having different perspectives about the issues. Quantitative analysis along with the qualitative study of the patterns of metaphor helped the author to differentiate all characters diachronically and synchronically.

Collocation and concordance are two concepts that are very close to frequency and in most cases, they appear together. Collocation is the sequence of the repeated co-occurrence of words within the corpus. As illustrated by Hori, in *The Bleak House* by Charles Dickens, the term "little" collocates with "scream", and as the extensive study of Hori showed, it indicates the behavior of a character and in this case, it indicates a cousin of Sir Leicester, Miss Volumnia.

1.6 Conclusion

This chapter reviewed the emergence, rise and development of digital humanities. It has been argued that critics define digital humanities as an attempt to putt different arenas of technology and criticism into contact with one another to form a new methodology of literary study. Digital humanities was addressed as an attempt to collect heterogenous phenomena under

a unified framework to examine the underlying meaning of the constituting elements. Also, four approaches within digital humanities were introduced that include the analysis of the digital tools, evaluation of the humanistic projects through the employment of digital tools, critical thinking about digital context and finally the multifacet humanities-driven and digital-driven approach. Then, Three temporal and conceptual phases of digital humanities were proposed that included early, current and future prospects. In the early phase, digital humanity was concerned with employing tools for linguistic purposes, along with answering the main questions within humanities such as how the transition from humanities computing to digital humanities took place and what kind of effects it had on the research arena. Current debates on digital humanities are mostly concerned with the adaptation of computational tools for scholarly examinations as well as discussion on moving from the traditional approach of focusing on the physical texts to the digital approach in which the print is no longer the core. Another significant emphasis of the current debates in digital humanities include the necessity of investing and forming digital contents and international initiatives that support the creation and management of digitalized contents worldwide. And the final phase include the prospects about the future of digital humanities that can go beyond the defined limitations of the science by employing algorithmic criticism.

The employment of digital tools in the analysis of literary texts was offered afterward. It was argued that the employment of the computational approaches to literature and language offer distant reading as the best methodology for quantitative and qualitative study of the literary texts and it supports the unraveling of the text's structure. In this regard, the best method of examing literary texts was offered as managing different degrees of distance and moving between close and distant readings.

Chapter Two

Diasporic Literature

> "Exile is strangely compelling to think about but terrible to experience"
> Edward Said, *Reflections on Exile and Other Essays*

2.1 Introduction

During the last quarter of the twentieth century, literary studies in particular and the studies of humanity, in general, moved into a new direction which completely opened a new field of inquiry, what is known as postcolonial studies. Since its inception in 1980s, postcolonial studies has grown tremendously in significance and currently postcolonial subjects form a very large section of research, publications, and teachings. Postcolonial studies is considered as an interdisciplinary field of inquiry where subjects as varied as literature, history, psychology, and politics intersect one another, which makes this field of inquiry a subject whose margins are rather vaguely defied and whose contours are difficult to grasp.

In the following, the most important terms that have come to occupy the center of contemporary literary discussions in relation to postcolonialism will be discussed and the relations between colonialism and the followings will be examined in detail: Western Culture and Colonial Discourse, Imperialism and its cultural effects on the colonizer and the colonized subjects and the Diasporic Discourse and Discursive resistance.

2.2 Western Culture and the Colonial Discourse

Any in-depth analysis of the cultural impacts of colonialism on the coloniser and colonised subjects requires an understanding of colonialism and its overall implications. According to Oxford English Dictionary, colonialism is "a settlement in a new country; a body of people who settle in a new locality, forming a community subject to or connected with their parent state; the community so formed, consisting of the original settlers and their descendants and successors, as long as the connection with the parent state is kept up" (*Oxford English Dictionary*, 76).

However, the definition provided by the Oxford Dictionary remarkably avoids any reference to people other than the colonisers and, there is a missing indication to the local inhabitants of the place prior to the establishment of colonies. The provided definition refers neither to the local inhabitants, nor to the confrontation of the two sides or even the domination of one side over the other. Besides, it neglects the fact that colonialism contains more than extracting tribute, goods and wealth from the countries it conquers; it restructures the economies of the conquered countries and draws them into a complex relationship with its own (*Colonialism/ Postcolonialism*, 9). Critics have also criticized the definition of colonialism, the scope it covers and its focal points, as well as raising questions about the aftermath of

colonialism referred as postcolonial era, in which the physical domination of a state over other states is outlawed.

One of the preliminary pitfalls of the postcolonialism as a concept is the vague notion of the prefix "post" which simultaneously indicates the temporal and ideological aspects of the term. Temporal in the sense of coming after or the "aftermath" of colonialism and ideological in the sense of replacing colonialism. In this respect, a country may be both postcolonial, for being formally independent and neocolonial for continuing to be economically or culturally dependent (McClintock, 87). Therefore, although in this perspective the new global arrangement does not embrace the direct domination as experienced earlier, it keeps the door open for economic, cultural and political dominations.

According to Iain McLean, the Post-colonial state, in its political meaning, refers to any of the new nation states that have come into being out of the decolonization process of the Second World War, also referred as the developmental state. The post-colonial state has two distinguishing features, including its political and economic agenda, and its capacity to rule. The post-colonial state has been characterized as strong or weak based on its capacity to implement political decisions, as well as the amount of available economic resources to that state (*Oxford Concise dictionary of Politics,* Iain McLean). Post-colonialism, which refers to the end of colonialism in its historical sense, came into existence by the end of World War II. The post-colonial and post-structural perspectives deconstruct the colonial discourse and question the meaning of the West all over again. This perspective highlights the significance of colonialism and imperialism on the emergence of the West and its perception of itself and the world around it. According to Ashcraft, this perception has facilitated the globalization of the western set of

values, identity and lifestyle and has moved the whole world into a post-colonial age. As Aschcraft and Ahluwaila in *Edward Said: The Paradox of Identity* declare:

> Colonial discourse theory is that theory which analyses the discourse of colonialism and colonization; which demonstrates the way in which such a discourse obscures the underlying political and material aims of colonization, and which points out the deep ambivalences of that discourse as well as the way in which it constructs both colonizing and colonized subjects. Post-colonial theory investigates, and develops propositions about the cultural and political impacts of European conquest upon colonized societies, and the nature of those societies' responses (32).

In general, post-colonialism can be referred as a collection of research on the cultural and political impacts of the colonialism on the colonized societies. As stated in the book *Edward Said*, the concept of "post" refers to the era "after colonialism began" rather than "after colonialism ended" because the cultural challenges between the imperial and dominated societies have never ceased (Aschcraft and Ahlowalia, 15).

2.3 Imperialism and the Cultural Effects on Coloniser and Colonised

Post-colonial theory is concerned with a range of cultural engagements, the impact of imperial languages upon colonized societies, the effects of Western 'master discourses' such as history and philosophy, the nature and the consequences of colonial education and the links between the Western knowledge and the colonial power. Postcolonialism is also concerned with the responses of the colonized; "the struggle to control self-representation, through the appropriation of dominant languages, discourses, and forms of narrative, the struggle over

representations of the place, history, race and ethnicity and the struggle to present a local reality to a global audience" (Ranjan Ghosh, 76).

Despite the differences among viewpoints toward the definition and the usage of the term Postcolonial, there is a possibility of defining a common ground among different narratives. From Yung's perspective, post-colonialism has emerged as an umbrella term for determining academic and multidisciplinary political, theoretical and historical studies which benefit from an international pattern and formulation. According to him, postcolonial studies does not seek to provide a new category as the former studies had done. Instead, postcolonial studies include academic endeavors attempting to determine what kind of misunderstanding and suppression the imperialistic hegemony has caused in the process of colonialism. Therefore, the postcolonial sociology is the result of dissatisfaction with the imperialist views of the colonised and attempts to create a basis for the study of society aside from Oriental assumptions (*White Methodologies*, 41-43).

Although postcolonialism has been previously employed to distinguish the pre and post-independence eras of given cultures which have been in close contact with colonialism, according to Aschroft, postcolonial literature is concerned with the cultures affected by the imperial process from the moment of colonization to the present day, mainly because there is a continuity of preoccupations throughout the historical process started by the European imperial aggression (*Empires Write Back*, 2).

2.4 Postcolonial Literature

The study of English Literature has primarily signified the study of British Literature, or at most, one can include the study of American literature, but in the new category of postcolonial

literature, a wide area of literary texts that come from different parts of the world as varied as India, West Indies, South America, Canada, Africa and Australia. Postcolonial studies is a fast expanding field of inquiry whose focal point is the relation of culture and colonial discourse that emerged by the end of the mostly Western colonial era and addresses the role of discourse and cultural factors in preserving colonial domination and encouraging to think through the backdrops of colonialism, resistance, cultural legacy of colonialism as well as decolonization as a process.

According to Quayson, more than focusing on the economic factors, it is the role of the cultural factors that is emphasized in this field of study. The postcolonial theory is a cultural approach that focuses on knowledge as a factor against Western societies, to emphasize not only the material factors behind the formation of colonial domination, but also the role of discourse and ideology to the extent that postcolonialism can be considered as a scientific field on how the knowledge as a cultural phenomenon is formed and criticized (*Postcolonialism: Theory, Practice or Process*, 2).

Postcolonial studies encompass critical investigation of culture, history, literature and discourses that shape the former colonies of the European imperial power, mostly focusing on the Third World countries located in Asia, Africa, Caribbean islands as well as South America. Edward Said in *Orientalism* revisited Michel Foucault's critique of the discourse for evaluating his theory of "cultural imperialism", by which he referred to the imposed imperialistic power not through force but by disseminating Eurocentric discourse that guarantees the dominance and supremacy of all that is related to "Occident", as opposed to the representation of the inferior "Orient" (2). According to Abrahams, there are several focal issues that the postcolonial studies revolve around including;

1. The rejection of the master-narrative of Western imperialism and its substitution by a counter-narrative, in which the colonial cultures fight their way back into a world history written by Europeans;

2. The formation of the colonial and postcolonial "subject" within the Western discursive practices and the features through which the subject conceives not only itself, but also the world around it;

3. Destabilizing the Eurocentric literary and artistic values and expanding the literary canons to include colonial and postcolonial writers (*A Glossary of Literary Terms*, 236-237).

From Said's point of view, imperialism includes the practice, the view and the perspectives of a dominant metropolis that decides about far lands. Said, in *Culture and Imperialism,* explained the features of imperialism and colonialism in details:

> Neither imperialism nor colonialism is a simple act of accumulation and acquisition. Both are supported and perhaps even impelled by impressive ideological formations that include notions that certain territories and people require and beseech domination, as well as forms of knowledge affiliated with domination: the vocabulary of classic nineteenth-century imperial culture is plentiful with words and concepts like "inferior" or "subject aces," "subordinate peoples," "dependency," "expansion," and "authority" (9).

Given the historical records, Edward Said believed that the European and American's interest in the Orient was primarily a political one. However, it was the culture that formed that interest along with political, economic, and military rationales that manifested the Orient as a complicated place. Edward Said explained orientalism as:

the distribution of geographical awareness into aesthetic, scholarly, economic, sociological, historical, and philological texts; it is an elaboration not only of a basic geographical distinction … but also of a whole series of "interests" which, by such means as scholarly discovery, philological reconstruction, psychological analysis, landscape and sociological description, it not only creates but also maintains; it is, rather than expresses, a certain will or intention to understand, in some cases to control, manipulate, even to incorporate, what is a manifestly different (or alternative and novel) world; it is, above all, a discourse that is by no means in direct, corresponding relationship with political power in the raw, but rather is produced and exists in an uneven exchange with various kinds of power (*Orientalism*,12).

Colonial studies lean on Foucault for its understanding of discourse. Foucault believed that there are defined structures that limit not only the formation, but also the circulation of discourse. It is for the purpose of institutionalizing the defined discourse that institutions including educational and legal systems were created. According to him, disciplines are "forms of power which are not violent or destructive, but rather seek to produce docile subjects through processes of training, correction, normalization and surveillance" (Gallagher, 53). Therefore, institutions are in charge of the overall control of the discourse through regulation and prioritizing and foregrounding certain discourses, while marginalizing and even confining certain other opposing discourses. Foucault in *Order of Discourse* defines the exclusionary process of discourse by saying that "in every society the production of discourse is at once controlled, selected, organized and redistributed by a certain number of procedures whose role is to ward off its powers and dangers and to gain mastery over its chance events" (Foucault, 210).

According to him, the formation and circulation of the desired discourse is underlined with its connection with power, and therefore; the kind of discourses that are prevalent in any given situation largely depends on the institutions which regulate and ratify the production and dissemination of knowledge. "Discourse is not simply that which translates the struggle or the system of domination, but is the thing for which and by which there is the struggle, discourse is the power which is to be seized" (Foucault, 52-53).

2.4.1 Oriental Point of View

The point that connects Foucault's ideology to postcolonialism is the connection he made between knowledge and power, that is to say, Foucault believed that it is the more powerful institution that gets the chance to regulate knowledge and its discursive manifestations. To put it differently, by highlighting the connection between power and knowledge, Foucault sought to convey this message that a concept turns into an area of investigation only under circumstances that the power has put as such, and therefore; the formation of knowledge and the methods of power unfold together and through one another. This relationship was more elaborately discussed in *Orientalism* published in 1978 by Edward Said. He is widely considered as the founder of the postcolonial studies and the critic of the colonial discourse, which he believed has been running in the history as well as literature. What Said attempted to connote is that the military and economic dominations of the West are tied with the discourse about the Orient. That is the same way that Foucault had linked power, knowledge and discursive manifestation of knowledge with one another, and in return, Said applied the general connotation of domination to the specific case study of the European domination over Orient. By confirming Foucault's idea that knowledge leads to power, Said also acknowledged that Western knowledge of the East

leads to Western world's control of the East. According to Ania Loomba, the profound

Foucauldian insight on the connection between power and knowledge has facilitated the

identification of the extent to which 'knowledge' about 'the Orient' was created and circulated as

an ideological accompaniment of the colonial power and nurtured and supported by other

disciplines including philology, history, anthropology, philosophy, archaeology and literature.

Loomba agrees that Edward Said had redefined the study of colonialism, but also believes that he

had underlined a certain type of perception and thinking associated with colonial power (60).

Said addressed the relations between the Orient and the Occident as a relation of power,

domination and a different degree of complex hegemonies and in respect to Oriental hegemony,

he further argued:

> The Orient was Orientalized not only because it was discovered to be "Oriental"
>
> in all those ways considered common….but the phenomenon of Orientalism deals
>
> principally, not with a correspondence between Orientalism and Orient, but with
>
> the internal consistency of Orientalism and its ideas about the Orient, despite or
>
> beyond any correspondence, or lack thereof, with a "real" Orient (*Orientalism*, 5).

Furthermore, the travel tales that were written at the time of the Renaissance were a

mixture of fiction, perspectives of the earlier times and the first-hand observations of what lied

outside the European boundaries separating inside and outside, the 'self' and the 'other'. The

outside worlds encountered by the European travelers were interpreted by them through

ideological filters, or ways of seeing, provided by their own cultures and societies. However, the

impetus to trade with, plunder and conquer these lands also provided a new and crucial

framework through which they would interpret other lands and peoples (Ania Lamba, 64). Yet,

Porter, claimed that Said flattened the "historical nuances into a fixed East versus West divide" (qtd. in *Colonialism/ Postcolonialism*, 46).

Also, by putting forward the fact that knowledge about the East was a part of the process of keeping the power over them, Said demystified the status of knowledge, by indicating the abruption of the distinction between ideology and objectivity (*Orientalism*, 45-46). For that reason, he rejected to accept the objective perspective of the West toward the East. One of the significant recurring critiques of Edward Said in respect to Oriental studies was that the Oriental point of view constantly suggests a binary opposition between East and West, which has been a more or less static feature of the Western discourses from classical Greece to the present day. He believed that this ideology run through all of the Western Literature and argued that Occident has represented Orient in a thousand ways, all of which add up to creating a version of East that is strange, exotic, sensual, barbaric, of a cruel place and the opposite of the rational just normal, civilized West. Said in *Orientalism* argues that in this perspective, Orient is "irrational, depraved (fallen), childlike, different"; thus, Occident is rational, virtuous, mature, and "normal". Yet, what gave the Oriental's world its intelligibility and identity was not the result of his own efforts but rather the whole complex series of knowledgeable manipulations by which the Orient was identified by West" (41).

Edward Said believed that whoever that conducts a research or teaches or writes about Orient should be referred as an Orientalist and what he does is considered as Orientalism. (*Orientalism,*2) He further argued :

> There are Westerners, and there are Orientals. The former dominates; the latter must be dominated, which usually means having their land occupied, their internal affairs rigidly controlled, their blood and treasure put at the disposal of

one or another Western power… Once again, knowledge of subject races or

Orientals is what makes their management easy and profitable; knowledge gives

power, more power requires more knowledge, and so on in an increasingly

profitable dialectic of information and control (36).

2.4.2 Discursive resistance

Along with Edward Said and Gayatri Spivak, Homi Bhabha is considered as a part of

"Holy Trinity" in postcolonial studies. Homi Bhabha's most prominent undertaking is a

collection of essays called *The Location of Culture* that was published in 1994. By crossing some

concepts such as race, class and gender, Bhabha widens his prospects of study and addresses

identity in the postmodern world as a conceptual issue which is formed and evaluated through

conflicts and differences. Like Edward Said, Bhabha examines the relationship between the

coloniser and the colonised, but his focus is not limited to the colonial power and discourse. To

make a balance in postcolonial studies, he evaluates the behavior, language and mentality of the

coloniser in relation to the colonised subject. For Bhabha, "the colonised subject is ontologically

incalculable" (*Postcolonial theory*, 135). Homi Bhabha further argues that the ambivalent

feedback of the colonised to the colonial power is "half acquiescent, half oppositional, always

untrustworthy—produces an unresolvable problem of cultural difference for the very address of

colonial cultural authority" (Bhabha, 33).

Unlike other post-colonial critics who address the colonised as an "other" who owns no

place to speak from and share his experience with coloniser, Homi Bhabha believes that there is

a space in which the colonised can communicate his experiences in form of concept to the

coloniser. In his rather different reading from Edward Said, he finds the relationship between the

colonised and the coloniser more complex and meaningful than what Said in *Orientalism* and Frantz Fanon in *The Wretched of the Earth* pointed to. Bhabha acknowledges the "force of writing and its rhetorical discourse as a productive matrix which defines the 'social' and makes it available as an objective of and for the action", but he argues that textuality is not just "a second-order ideological expression or a verbal symptom of a pre-given political subject" (*The Location of Culture,* 22-23).

From Bhabha's point of view, the identity and the personality of individuals in the colonial contexts are heavily influenced by the unconsciousness of both sides. Both the coloniser and the colonised are affected by their unconsciousness in their actions. Therefore, the coloniser is not immune, neither physically nor politically as Said believed. In fact, in his perspective, the boundaries between the white and black, the self and the other get smaller and the previous background for the preservation of racial identity in its traditional form disappears. By referring to Fanon's *black Skin White Mask*, Homi Bhabha examines the relations between the coloniser and the colonised. In his review on this book, Bhabha mentions that remembering is an inevitable yet dangerous bridge between colonialism and cultural issues. In *The location of Culture*, he writes:

> Remembering is never a quiet act of introspection or retrospection. It is a painful remembering, a putting together of the dismembered past to make sense of the trauma of the present. It is such a memory of the history of race and racism, colonialism and the question of cultural identity, that Fanon reveals with greater profundity and poetry than any other writer" (90).

Edward Said, in *Orientalism* focused on the military, legal, economic and political relations of the coloniser and the colonised, while Homi Bhabha dealt with the cognitive

relationship between the dominant and subaltern cultures. From his point of view, in the colonial relations, the identity of neither side is genuine and both sides need each other to build their identities. In fact, instead of examining the relationship of the dominant and the subordinate through political approach, Homi Bhabha benefits from the psychological approach. He believes that cultural differences establish a dialogue between the colonised and the coloniser. He is affected by Jacque Lacan and claims that the colonised subject will be able to resist the colonial power once he manages to gaze back at the coloniser and by doing so, he will challenge the coloniser. The colonised accomplishes this goal through mimicry. In fact, the political imitation is a strategic method that can help the defeated to dominate. This form of resistance is somewhat self-conscious and somewhat unconscious. Accordingly, he believes that the colonial culture is not absolute in terms of power and as much as it affects, it can be affected. Therefore, as much as mimicry is the tool of power, so is a tool of resistance. This imitation and the changes the colonised causes in the culture of the coloniser continues to the extent that it ultimately leads to a balance between the dominant and the subordinate (Gilbert, 72-75).

The term "mimicry" has been crucial to Bhabha's view of the ambivalence of colonial discourse, because as much as the mimicry is the process by which the colonised subject is reproduced as "almost the same, but not quite" (Bhabha, 86), it can be equally a threat for that discourse, for copying the coloniser's culture, behaviors, manners and values by the colonised contains both mockery and a certain 'menace', "so that mimicry is at once resemblance and menace" (86).

2.5 Diasporic Studies

While Oriental studies focus on West's understanding of the East, the focal point of the Diasporic studies is the experience and the reading of the Orient of the Occident. Therefore, Diasporic literature forms an integral part of the broader category of postcolonial literature. Diaspora is a concept rooted in the dispersion or exile of the Jews from their homeland. Originating from the idea of displacement from a homeland, diaspora points to those "communities of people who have been dislocated from their native homeland through the movements of migration, immigration, or exile" (Braziel and Mannur, 1). Kevin Kenny in *Diaspora: A very Short Introduction* investigates the origin of diaspora by saying:

> The Greek noun diaspora derives from the verb diaspeirein, a compound of "dia" (over or through) and "speirein" (to scatter or sow). The word emerged from the proto-Indo-European root, spr, which can be found todady in such English words as "spore," "sperm," "spread," and "disperse." In all of its various uses, diaspora has something to do with scattering and dispersal. To the ancient Greeks, diaspora seems to have signified mainly a process of destruction. Epicurus used diaspora to refer to the decomposition of matter and its dissolution into smaller parts. Human communities subject to the destructive force of diaspora were similarly split asunder. Thucydides employed diaspora in this way, in a minor passage in the History of the Peloponnesian War, to describe the Athenians' destruction of Aegina and the banishment and dispersal of its people (2).

Oxford English Dictionary defines diaspora as the dispersal of the seeds as well as the migration of people from one place to another (Oxford English 8 Dictionary, 1989). Therefore, it

could be argued that diaspora, beside connoting the notion of a center, addresses a point of departure and a journey to put roots somewhere else. Bill Ashcroft et al. considered diaspora as a historical fact and describe it as a "voluntary or forcible movement of people from their homeland into new regions. They consider colonialism as the diasporic movement, dispersion and settlement of Europeans all over the world" (68). According to Arnold Ages, in regard to the Jewish people, diaspora refers to different eras in the history of Jewish community, representing their status during the Babylonian captivity in the sixth century B.C, the flourish of the Jewish community of the Alexandria prior to the rise of Christianity, the revolt against the occupation of Palestine in the first century by Romans as well as the destruction of the temple at Jerusalem in 70 A.D., which all led to the dispersion of the Jewish community worldwide (3).

However, critics came to agree on the more general meaning of diaspora referring to the scattering of people who are connected by a sense of homeland, imaginary or otherwise ("The Turn to Diaspora", 12). According to Baumann, diaspora is a derivation from a Greek term meaning "to scatter", "to spend" and "to disperse" and it is emotionally charged with "uprootedness, precariousness and homesickness" (Martin Baumann, 314). Diasporic journeys are about "settling down and putting roots elsewhere" (Brah, 182). Rogers Brubaker refers to the diasporic community as a minority group with an imaginary or real homeland orientation that hold on the collective memory of the ideal homeland and carry three main features; 1) dispersion in space 2) orientation toward homeland and 3) maintaining boundaries ("The Diaspora's diaspora", 5).

According to James Thomas Zebroski, the capitalized form of diaspora refers to the Jewish people living outside Palestine after the Babylonian exile ("Theory in the Diaspora", 676), the type which is also referred as "the ideal type" by Webster (qtd. in *Global Diasporas: An Introduction,* 5).

2.5.1 Diasporic Literature

Diaspora has found its way into academic arena in the last decades and the interest in diasporic has intensified since then. Diaspora offers the possibility of examining self, identity, space and belonging in relation to one another, but not as fixed ideologies but in the process of becoming. Diaspora provides the possibility for a non-essentialized self, therefore, it interrupts the established relationship between the place and identity by keeping multiple connections between the present 'here', and a past 'there'. Accordingly, Clifford supported the fact that "dwelling *here* assumes a solidarity and connection *there*" (296). The relation between homeland and away connotes the diasporic understanding of the term which is associated with a sense of loss. Diaspora, in this sense, includes a fundamental idea of a fatherland that is scattered.

According to Floya Anthias "… The original father (land) is a point of reference for the diaspora notion: it is this constant reference point that slides into primordiality" (qtd. in *Diaspora and Hybridity*, 16). However, the contemporary diaspora seems to pose a third space, a place which is "beyond space and time, and beyond the situated practices of place and the lived experience of history" (Mitchell, 534). In this sense, diaspora has been transformed from referring to the condition of Jews in exile, to covering ethnic, religious and national communities who are living outside the place they were originally located in. Khachig Toloyan, in his introduction to the very first edition of Diaspora, wrote, "Diasporas are the exemplary communities of the transnational moment" (3). As Brubaker noted, the strand of the diasporic literature is "firmly rooted in a conceptual homeland" (2). Homeland, either real or imagined, is described as an authoritative source of value, identity, and loyalty, and diasporas are described descriptively with reference to that origin. Robin Cohen has divided diasporic communities into five categories, all of which have their own defining features as well as similarities that include

44

Victim diaspora, Worker diaspora, Commercial diaspora, Imperial diaspora and the Cultural diaspora (*Global Diaspora*, ix-x). By addressing the current diasporic communities such as Armenian, Turkish, Palestinian and so on, Safran underlined the Jewish diaspora as the "ideal type" (84).

Misti L Williams argues that the myth of return to the homeland is a crucial factor in the diasporic consciousness, for assimilation to the host country seems impossible and the eventual return is desirable. However, this return is not necessarily a physical return or a plan to return in the future, but the return ideology makes the experience tolerable (10). As stated by Safran "the myth of return, serves to solidify ethnic consciousness and solidarity when religion can no longer do so, when the cohesiveness of the local community is loosened, and when the family is threatened with disintegration" (91). However, according to Hall, the problem is that members of diaspora carry a trace of their particular language, culture, histories and tradition that has formed them into the person they are, but the problem is that they will not be able to be unified in the old sense, because they have been affected by several interlocking cultures, and cultures, belong at one and the same time to several homes ("Cultural Identity and Diaspora" , 310). Stuart Hall has advanced two ways of thinking about cultural identity in the context of defining diaspora. The first position defines it in terms of a shared culture; a sort of collective "one true self" that "reflects the common historical experiences and shared cultural codes which provide us, as 'one people,' with stable, unchanging and continuous frames of reference and meaning, beneath the shifting division and vicissitudes of our history." The second position of cultural identity is a matter of becoming as well as being: the significant difference which constitutes what we really are or what we have become (*Cultural Identity and Diaspora*, 225).

This situation creates a condition in which members of diaspora are neither living here, nor there and they are forced to be in the in-between space, where according to Bhabha is the "cutting edge of translation and negotiation" (*The Location of Culture*, 38) that he puts as the third space. Believing that any attempt for the cultural translation leads to ignorance of the essentialism of a prior given original culture, Homi Bhabha conceives hybridity as the aftermath of difference and the attempt to translate and he states:

> But for me, the importance of hybridity is not to be able to trace the two original moments from which the third emerges; r hybridity to me is rather the 'third space' which enables other positions to emerge. This third space displaces the histories that constitute it, and sets up new structures of authority, new political initiatives, which are inadequately understood through received wisdom… (Hybridity is) just like a translation, so that hybridity puts together the traces of certain other meaning or discourses. It does not give them the authority of being prior in the sense of being original: they are prior only in the sense of being interior the process of being anterior. The process of cultural hybridity gives rise to something different, something new, and unrecognizable, a new area of negotiation of meaning and representation (211).

Ankie Hoogvelt in *The New Political Economy of Development* positions hybridity at the center of both colonial and postcolonial discourses, but in the opposite directions. Hoogvelt claims that in the colonial era, hybridity was a debilitating term in reference to the lowest form of human life as well as mixed breeds who were white but not completely. However, in postcolonial discourse, hybridity is both celebrated and privileged as a type of higher cultural intelligence as the outcome of the in-between space of two cultures that enables the negotiation

of the existing differences (158-159). According to Papastergiadis, Homi Bhabha developed his understanding of hybridity from both cultural and literary theories to address the formation of identity and culture under colonial antagonism and inequality. Meredith argues that for Bhabha, hybridity is a process in which the colonizer is committed to translate the identity of the colonized Other in the single universal framework, but then he fails to produce something familiar and instead he forms something new (2). As stated in "Together-in-difference: Beyond Diaspora, into Hybridity", the relationship between cultural theorists and hybridity is as follows:

> For postcolonial cultural theorists such as Stuart Hall, Paul Gilroy, Trinh Minhha, Homi Bhabha and others, hybridity has an explicitly critical political purchase. They see the hybrid as a critical force that undermines or subverts, from inside out, dominant formations through the interstitial insinuation of the "different", the "other" or the "marginalised" into the very fabric of the dominant. …The politics of hybridity here then is one of the active interventions, involving both a disarticulation of exclusionary conceptions of "white" and its rearticulation as a necessarily impure and plural formation that can no longer suppress the black other within. In this sense, hybridity … destabilises established cultural power relations between white and black, coloniser and colonised, centre and periphery, the "West" and the "rest", not through a mere inversion of these hierarchical dualisms, but by throwing into question these very binaries through a process of boundary-blurring transculturation (9).

In an oft-quoted remark by Stuart Hall, diasporic experience is defined "… not by essence or purity, but by the recognition of a necessary heterogeneity and diversity; by a conception of 'identity' which lives with and through, not despite, difference; by hybridity"

(Hall, 235). Robin Cohen has divided the progress of the concept diaspora into four historical periods, each focusing on the characteristics of diaspora in that era. From 1960s to 1970s, Diaspora which was solely used with a capital initial and in a singular form contained the traumatic experiences of African, Armenians as well as Irish people. By relying on the observations of Safran and Tololyan, Cohen addressed the second stage of the evolution of the diaspora in 1980s when diaspora was employed to refer to the different groups of people as wide as political refugees, immigrants, ethnic minorities and expellees. One feature they all had in common though, was their either forceful or voluntary migration from their homeland. The third phase according to Cohen initiated in 1990s and included a reconsideration of the two main diasporic ideologies that included homeland and ethnic community. Given the complexities imposed by the postmodern ideologies, reconsideration of the definition of diaspora seemed inevitable. The final stage of diaspora refers to the contemporary time addressing the consolidation of theories, as well as a partial return to the tendencies of the first stage. According to Cohen, homeland and the desire to connect to the origin is still dominant. However, due to the acceleration of globalization process, concepts such as nation and border have undergone significant changes and will continue in this direction (*Global Diaspora: An Introduction,* 2-4).

2.6 Conclusion

This chapter examined the main characteristics of the postcolonial studies, its principal thinkers as well as its role in the emergence of the Diasporic literature. This chapter also touched on the most significant terms that came to occupy the center of the postcolonial discussions, including colonial discourse, imperialism and discursive resistance. It was argued that the focus of the postcolonial studies has been on the relation of the culture and colonial discourse, along

with evaluating the role of discourse and cultural factors in preserving the colonial domination. In addition, the postcolonial studies seek to address the formation of the colonial and postcolonial subjects, reject the master-narrative discourse, and destabilize the western-oriented literary values to include the works of authors from the East as well.

Edward Said was introduced as one of the most significant postcolonial thinkers who questioned the connection of power and knowledge and the ways it lead to the domination of one over the other. It was stated that while Oriental studies focuses on the West's understanding of the East, the focal point of the Diasporic studies is the experience and the reading of the Orient from the Occident. Therefore, Diasporic literature was pointed as an integral part of the broader category of the postcolonial literature. It has been argued that the meaning of diaspora has transformed from only referring to the dislocation of the Jews to include communities of people who have been dislocated from their native homeland through the movements of migration, immigration, or exile. Significant topics in diaspora were also mentioned that included living in-between, being marginalized, the myth of the eventual return to the homeland, and hybrid identities.

Chapter Three

South Asian American Diasporic Literature

3.1 Introduction

In chapter two the main features of postcolonial studies and its connection with the diasporic literature were examined. The main concepts that have occupied the main discussions of the diasporic literature were addressed. It was mentioned that in this ideology, knowledge is considered as power and the Western knowledge of the East has turned into the dominance of the West over the East. In this chapter South Asian Americans are introduced as one of the most significant representative of the East and the discursive dominance of the western conception on them will be evaluated. Also, the settlement process of the South Asians within the United States will be addressed and the main challenges they have faced along the way will be discussed. Their diasporic literature will be examined to identify the portrayal of identity, home, and the host country.

3.2 Migration of South Asians to the United States

The 1965 Immigration Act can be considered as a fundamental change in the immigration process of the United states that overturned the restrictions that had been in place since 1920s based on origin quotas and underlined the significance of family reunification. This legal amendment not only resulted in an uptrend immigration in general, but also altered the trend from the mostly European countries of origins including England, Germany and Ireland to Asia and Latin America. According to Timothy J Hatton, the 1965 Act was a defining moment that changed the American demographic figure entirely and established the basis for everything that followed. By reversing the half-century legal discrimination as Maira puts it, South Asians constituted the fastest growing minority group in the United States (Saran,71). Indians formed the largest part of the South Asian community, followed by Pakistanis, Bangladeshis, Nepalis, Sri Lankans, Bhutanese, and Maldivians (Srirama,17).

Several studies have categorized phases of migration from South Asia to the United States highlighting the diverse socioeconomic profile and class standing of the migrants. For instance, Rahma Zayna divides the arrival of South Asians into three major categories. He explains that the first group of South Asians arrived between 1897 and 1924, made up of mainly illiterate male Sikh and Muslim peasants who were fulfilling cheap labor needs and were treated with much racism. Unlike the first phase that mostly included Indian, the second wave of migration was drawn from all over India, as well as Pakistan, Bangladesh, Sri Lanka and Nepal. The 1965 Immigration law, which was introduced to accommodate labor market requirements, facilitated the arrival of middle class professionals, college educated and those who were seeking advanced training. The third wave of immigration, which brought compelling dramatic shift and

52

polarization in the South Asian community happened during the1980s. It was constituted of the established South Asian Americans who were sponsoring their families through the Family Reunification Act and diversity visas, who were in return less educated and less fluent in English and were driven toward blue-collar jobs (34-36).

Different reasons behind the voluntary or forceful migration as well as the social and economic status of migrants are among the critical factors for determining the depth of their diasporic experience as well as the intensity and speed of their assimilation in the host country. In that respect, South Asian Americans vary considerably based on the economic, social, ethnic, linguistic and religious make-ups (Sunaina Maira, 53). However, South Asians, like their Asian American peers, have been described as the model minority by the dominant voices in the United States to reinforce the myth of American meritocracy ("(Re)Telling traditions", 3) and accordingly, a "model minority stereotype" of them was presented to portray them as a homogenous ethnic group and excluded those that did not fit in the minority stereotype (Amita Shah, 3). The introduction of this framework into public discourse and literature led to discussions on the Oriental perspectives of the United States that attempted to present the civilized superior United States against the uncivilized and backward others and here, south Asians. This ideology is identified with the predominance of the white cultural discourse over the embodied other. This perspective, as proposed by Edward Said in his *Orientalism,* argues that the Western countries including America hold a subconscious perception of the East as inferior and uncivilized.

3.3 Oriental Perspective toward South Asians

Said puts Orientalism as a " Western style for dominating, restructuring, and having authority over the Orient" (3) and argues that historically, colonizers managed to employ dominant discourses, power structures and social hierarchies to present a fixed representation of the colonized other. By doing so, Said discusses that the West has succeeded to form an image of the East that is inferior, backward and irrational and in return, presented the West as superior, rational and progressive (Macfie, 8). Therefore, in the same manner that the perception of the West as rational and progressive was not possible without the portrayal of the East as inferior, the considered stereotypes for South Asian Americans have highlighted the superiority of the American culture and identity. American Orientalism puts bodies in an ideological hierarchy in which they are forced to indicate their belonging to either/or category and no place in between is allowed. Once the "strict relationship of … us and them is established", they should be translated and evaluated upon that and those who do not belong to the binary layout, then "will be excluded even from simple exclusion" (Jain, 2). Edward Said puts Orientalism as a politically charged system of beliefs and representation in which the hegemonic West leads Orient to its collective consciousness and preexisting culture.

Oriental representation is embedded in the 19th century Western colonization and has formed a powerful dichotomy and by doing so, has established its own dominance over the East. This ideology is a deep basis for evaluating the method that the West employs to represent the East as well as the South Asian identity (91). Homi Bhabha, in *Nation and Narration,* argues that within this Oriental perspective that the dichotomy between the East and West is preserved, the hybrid identities are created in which people feel to be in-between ideologies and cultures. This position is what Bhabha refers to as the Third Space that allows other positions to emerge by blurring the limitations of existing categories and calling attention to fixed categories of culture and identity. Chaudry

defines hybridity as "the fluid state of having multiple, shifting identities which are constructed and differentially privileged in response to contextual demands for alienation and allegiance." (48) In the same manner, migration and the hybrid identities of the South Asians have always been and will still continue to be the outcome of the "shifting imperialisms and neoliberal globalization" (Bald et al, 7). In this situation, immigrants have to define everything around the dominant social group within which they function. Therefore, by default and by necessity, they are forced to define themselves against the dominant group.

Given that migrants are forced to negotiate their identities within the dominant ideology regularly, the process of developing unique identities when the ethnic minority stereotypes are also associated to them become more complicated. It is also the same about South Asian Americans. As Pang puts it, despite their varied differences, South Asian Americans are still perceived by the mainstream society as Asian Americans that all look, think and act similar (Pang, 379). According to Tamara Bhalla, the discourse of the South Asian American community is the byproduct of the liberal multiculturalist ideology that portrays Asian Americans as a model minority and relies on the myth that the United States is a color-blind state that rewards the hard work, education as well as the class standing that seem to be the characteristics of certain racial groups (26).

However, categorizing Asian Americans based on race has had negative implications for them as well. Teranishi and Tchen argue that race for Asian Americans is "flawed, highly politicized classification system that is constantly shifting along with the American landscape" (15). Therefore, they have been expected to meet a certain defined social and political agenda that makes it harder to fit in because race has an unstable and flowing meaning, and racial difference is viewed as being racially superior or inferior to one another. Thus, immigrants from South Asia are forced to face the portrayal of their identities as other and inferior compared to that of the United States (Amitah Shah,

35). According to Shah, South Asian Americans establish their identities based on how they are perceived by the society, how they see themselves as well as how their different life experience has shaped their sense of beings (91). Kuan Chen, in *Asia as Method,* addresses South Asians as "a critical proposition to create new possibilities and to transform the existing knowledge structure and to advance the comprehensive decolonization process" (XV). Identification of the South Asian Americans as a minority group carries two main rationales: first, this classification puts South Asian Americans in contact with the more settled pan-ethnic formation of Asian Americans and second, it corrects the accepted hegemonic ideology that the region is only India and Hindu-centric (Tamara Bhalla, 9). The emergence of the Asian American discourses and Asian Studies aimed at evaluating the assimilation of Asian and Asian American identities in the global context and to examine the hybridized East Asian Subjects in Diaspora along with narrating the untold stories of marginalized people.

3.4 The Diaspora of South Asians in America

Once evaluating the South Asian American literature, it is important to study the progress that Diaspora, as a critical method of study, has gone through to identify the present situation of it within South Asian American community. According to Khachig Tölölyan, the portrayal of the migration process has been totally changed by the historical events and he explains "The term that once described Jewish, Greek and Armenian dispersions, now shares meanings with a larger semantic domain that includes words like immigrant, expatriate, refugee, guest-worker, exile community, overseas community and ethnic community" (4-5). Diaspora originally connoted dispersion or scattering of the people, specifically the exile of the Jews from their historical

homeland, along with the oppressions they endured as a result of dispersion. Diaspora could also connote the dispersion of the people, culture or language that was formerly concentrated in one place. According to Mohanram, diaspora refers to those who are "perceived as being out of place, either from their natural environment or their national boundaries" (25). Safran addresses a group of critical features regarding diaspora, including the idea of an original center, a collective myth of the homeland, a sense of alienation in the host land as well as the myth of the eventual return (XIV).

Diaspora not only indicates a sense of exile and trauma within Asian Americans' migration and settlement, but also shows a progressive generation of new identities which settle outside their homeland to be able to cope with living inside, with a difference (Divya Girishkumar, ii). Exile, within diasporic experience, as Said in *Reflections on Exile and Other Essays* put is:

> strangely compelling to think about but terrible to experience. It is the unhealable rift forced between a human being and a native place, between the self and its true home: its essential sadness can never be surmounted. And while it is true that literature and history contain heroic, romantic, glorious, even triumphant episodes in an exile's life, these are no more than efforts meant to overcome the crippling sorrow of estrangement (173).

Diaspora is a social construct in which different elements including feeling, consciousness, memory, shared identity, history and narratives work together to establish a reality in which a sense of association with a homeland, either real or imaginary, opposes both forgetting the past and assimilating in the current space and time. As stated by Clifford, diaspora can be considered as a "loosely coherent, adaptive constellation of responses to dwelling-in-displacement" (Clifford, 287). and it can be identified as a signifier "not simply of transnationality and movement, but of the

57

political struggles to define the local, as the distinctive community, in historical contexts of displacement" (Clifford, 315). Within this framework, home is not only a physical entity referring to a geographical location or settlement, but it turns into the ideal center for developing imaginative conceptualization of home as well.

Given that much of the exile life is occupied by a disorienting loss, caused by the sense of being an outsider in the new world, Edward Said has underlined both the pain that those in exile are forced to carry and their hesitation to belong to the new world and he writes:

> No matter how well they may do, exiles are always eccentrics who feel their difference (even as they frequently exploit it) as a kind of orphanhood. Anyone who is really homeless regards the habit of seeing estrangement in everything modern as an affectation, a display of modish attitudes. Clutching difference like a weapon to be used with stiffened will, the exile jealously insists on his or her right to refuse to belong (*Reflections On Exile: And Other Literary And Cultural Essays*, 183).

In addition, diaspora is considered a deeply conservative idea ingrained in narratives of exile from homeland. It is also a state of disruptive impurity that pictures identities and cultures in their post-ethnic emerging situation (Cheyette, xiii), along with carrying a sense of fragmentation within cultural identity and belonging and instead provoking a longing for home. Butler has identified four main features for diaspora that include: experiencing two destinations or more upon dispersal, having connection to either an actual or imagined homeland, having group identity and finally, experiencing a historical and temporal period for at least two generations (qtd. in *Diasporas and Interculturalism in Asian Performing Arts*, 4). Other scholars, such as Walter Connor, have taken a neutral implication of diaspora into account, only referring to "that segment of people living outside the homeland" (qtd. in *Global Indian*

Diasporas: Exploring Trajectories of Migration and Theory, 31). Because of the impossibility of returning home for an indefinite time, other critics, including Clifford, address diaspora as an experience of exile. The concept of return has undergone a significant change as much as the concept of diaspora itself has. The diasporic community who considered the physical return to homeland as the ultimate way to heal their fragmented identities, either find it impossible to return or were reluctant to return when it was possible. Therefore, over time, returning to motherland and the root of identity became possible through the imaginary and emotional return. Also, unlike the first generation of South Asian authors who narrated the bitter realities of the diasporic life and underlined the significance of the eventual return, the second and the third generation of writers provide a detailed and authentic map of the existing realities that have turned the desire for a return into an impossible mission.

It is possible to examine diasporic literature by taking some of its key characteristics into account. First, it is based on the idea of a homeland from where the displacement takes place. Second, the diasporic literature narrates the journeys undertaken for different reasons. In addition, it provides accounts of another "sense o place", a place away from homeland and finally, it could read how the protagonist behaves in the new place either by adopting or rejecting the new cultural codes (Pundir,70). By relying on S.Vertovec's theories of diaspora, Yasmin Hussain, in *Writing Diaspora: South Asian Women, Culture and Ethnicity,* provides a different interpretation of diaspora. For examining diasporic literature as well as diasporic experience, Yasmin Hussain puts diaspora on four different categories: 1) Social category including the forced displacement along with the negative impacts of alienation and loss, 2) Individuals in a diasporic community, connection to a place of origin and the sense of being far from it and sharing that feeling with those who have the common root, 3) Diaspora as a mode of cultural

production expressed in a hybrid form of expression due to an ongoing process of formation and reformation, 4) diaspora as a problem and immigrants as a threat to the security and integrity of the host country (6).

3.5 South Asian American Diasporic Literature

Traditionally, Asian American literature was considered as a group of texts written in English that portrayed the special social history within the life of individuals of different ethnicities who had faced discrimination based on laws and understandings that framed them as aliens. South Asian American authors have also contributed to the representation of the South Asian community in the United States as well as providing a deeper understanding of their diasporic experiences and the challenges and struggles they have faced both in terms of how they are perceived in the host country as well as how they perceive themselves. According to Bald et al. the history of the integration of South Asians in America can be divided in three eras encompassing 1917 to 1924, after 1965 and post 9/11. During the first era, South Asians were mostly blocked from entering the United Stated because of a series of law and regulations that basically addressed them as an ineligible racial group for citizenship; the second era, however, included naturalization of a larger but selected group of either highly educated or highly skilled migrants from South Asia; and the final era included the "War on Terror" era and its afterwards, in which South Asians and Muslim migrants were discriminated for surveillance, incarceration and deportation (3- 4).

Layfield divides the literary texts that have been written by Asian American writers into two aesthetical and formal groups, consisting of memoirs or autobiographies as well as ethno-racial bildungsromans. The main reason for the dominance of these two groups lies within the full potential of these types of narration that facilitate the intimate narration of stories in either fictional or

nonfictional settings (64). Several main subjects can be identified within South Asian American

diasporic literature. According to Bhatia, diasporic literature can be identified by examining whether

the characters are looking for a connection to their homeland and whether they tend to rediscover

and create new identities in the host land, all of which can cause conflicts, a sense of isolation,

belonging and so on. The idea of going on a journey is also a prominent theme in the diasporic

literature that brings the challenges along the way. In that respect, diasporic journeys are "about

traveling elsewhere and putting down roots. They are about arriving at a new location, settling down,

and having a memory and a longing for 'elsewhere' or another place" (Bhatia, 79).

The concept of home is also another significant subject that many of the diasporic literary

texts are revolving around and the authentic picture of the diasporic lives is only presentable in

the shadow of the connection of the identity to home. According to Divya Girishkumar, home is

not a psychological feeling, but it is a powerful sensation where a number of social, economic

and political elements contribute to the making of a home (3). In this sense, the narrative of the

lost homeland and memories of it, accommodate the diasporic communities to construct an

imagined homeland which is idealized, yet not possible to return to.

3.5.1 Women and South Asian American Fiction

Women have played a significant role in developing the position of South Asian

Diasporic literature in the United States. However, it should be noted that the role of South Asian

women in literature should be divided into the role of female authors and critics, as well as the

representation of women within the suggested frameworks. South Asian women are often

stereotyped to be permissive women who are forced to abandon their individual desires and

dreams for the sake of family's well-being. Many studies have examined the subordinate role of

women, the arranged marriage and in general, the very choice of women to decide for themselves. The significant stereotypes about South Asian women are focused on their femininity and sexuality as well: oppressed yet exoticized women, controlling men, arranged marriages, Indian cuisine as a signifier of domesticity, and female feticide (Tamara Bhalla ,29). According to Puar, these assumptions lean toward a perspective that South Asian women are built in a strict, inflexible, patriarchal and extended south Asian family in which the westernized second generations are obliged to stand against their own culture to set themselves free (qtd. in "Third World Woman", 84).

In the introduction to a collection of essays entitled as *A Patchwork Shawl*, Das Dasgupta addresses the lives of South Asian women including herself as "diverse and different, yet tied together with a common thread: experiences we have shared as immigrants from South Asia and women of color living in the United States of America" (1). South Asian women also have the potential to be discriminated against on the basis of their gender. For example, stereotypes of South Asian women suggest that they are passive, submissive, illiterate, and they smell like curry (Bannerji, 1993; Samuel, 2004).

The perspective of American society based on the existing stereotypes in regard to South Asian women is also criticized by different authors. As Shamita Das Dasgupta claims the general populace still tends to perceive South Asian Women based on the stereotypes as "docile, subservient, passive, politically unaware, asexual, and bound by tradition". She also argues that the Western conception refuses to accept any changes in its perception of the South Asian women and it only gives options to choose between two roles "traditional (read: backward, oppressive, sexist, uneducated, passive and docile) and progressive (read: enlightened, egalitarian, active, vocal, and Westernized). The pervasiveness of these dichotomous images has

obliterated any other viable models from our minds. Such linear and dualistic concepts are, of course, typical of Western thinking" (*A Patchwork Shawl: Chronicles of South Asian Women in America,* 11-12). However, many second-generation South Asian authors have refused these imposed stereotypes and roles, claiming other ways of living and being, by trying to define themselves in new forms. By avoiding to be "too ethnic" or " too American", the second generation of the South Asian Americans try to find harmony in developing their authentic identities (qtd in "Exploring Hybrid Identities", 93) and by doing so, they neglect to let others define them and they revisit the lines that had been drawn for them by the South Asian society, parents, as well as the prevailing society and instead, they come to understand their conflicting identities (qtd. In "Exploring Hybrid Identities", 20).

According to Yasmin Hussain, South Asian female authors are either indigenous people from India, Pakistan, Bangladesh and Sri Lanka or they are of South Asian descent but dwell in diaspora (53). The contributions of these women in South Asian American Diasporic literature have given access to the depth of the cultural and traditional issues within this community and by shedding light on the complex structure of them have provided new understandings that were not possible otherwise. The fiction that these women write differs from that of their male counterparts in several ways, including their depiction of the female characters. Also, the author's own definition of femininity could be observed within the portrayal of the female protagonist and its representation of the new woman (Uzma Akhand Hossain, 20). In addition, many studies have been conducted in the long-term absence of the female authors within south Asian literature, along with raising questions about male authors' supremacy especially that of Salman Rushdie and Hanif Kureishi (Divya Girishkumar,158). Gradually, South Asian female authors have provided counternarratives through their fictions and by doing so, they have turned the textual configuration of the women into a

political agent and "its discourse has developed as real and historically significant as she actually was" (Hussain, 53).

3.6 Critical Approaches toward South Asian Diasporic Literature

In general, the approaches that have been taken by the critics, scholars and authors about the representation of the South Asian diasporic literature can be divided into three main categories:

3.6.1 East Versus West Approach

By focusing on the Oriental perspective of the West toward East, the first approach underlines the employed hegemony by the United States to portray itself as superior and the minority ethnic groups, here, South Asians, as inferior. This approach focuses on the doctrines of Edward Said and other critics to criticize the ways in which South Asians have been employed to strengthen American discourse and authenticity. In this confrontation, that is mostly defined based on the power hegemony, immigrants are treated as Others who belong to periphery and their behavior are "almost the same, but not quite" in comparison to those in center. Meanwhile, diaspora is considered as a third-space in which the hybrid identities of those in exile are employed to highlight the American Discourse through underlying the stereotypes that are associated with them.

There have been extensive studies on the emergence and development of stereotypes regarding different model minorities in the United States which stem from an ideology that sought to approve the American meritocracy. These studies have shown that model minority stereotypes tend to convey this message that the political structure of the United States' society

facilitates the fulfillment of the American Dream for all. In this sense, Asian Americans and South Asian American in specific, have been portrayed as the role model for other minorities to follow. However, the failure of any minority groups is pictured as a consequence of the lack of industries and values rather than that of the domination of the Oriental perspective (Bic Ngo, 418). Therefore, it is safe to say that stereotypes were employed to contain and control minorities. In terms of the diasporic literature, one of the dominant stereotypes regarding South Asians has been the portrayal of South Asianness as a quality of an Other that desires for stereotypical Americanization that also confirms the oppositional positions of South Asians and Americans (Naber, 89). In this regard, the notions of inside and outside have also played a significant role in determining the discourse of nation, beauty, culture and so on. In oriental discourse of the United States, South Asians are depicted as backward, inferior and outsiders and in return, some parts of the South Asian community have embraced and carried this stereotype to distinguish themselves from blacks, Latinos as well as whites, by claiming their culture and tradition as superior (Prashad 2000).

Superiority of men over women along with the cultural gender expectations are also other significant stereotypes regarding South Asians that many authors have written about. For instance, Shamita Dasgupta refers to this fact that gender is at the center of any search for identity and that quest is harder for women because they should struggle to define themselves inside the conflicting cultures and traditions they are exposed to and at the same time, they should confront the stereotypes that have already defined them (98).

There is another perspective closely associated with this approach and that is considering the United States as the land of justice and the Americans as the savior who "save brown women

from brown men" and give voice to the suppressed people (Spivak, 296). In this framework, South Asia is the ultimate failure and the United States is the only answer to the problems.

3.6.2 Self-Critical Approach

In this approach, critics and writers challenge their intellectual and cultural heritage and evaluate them thoroughly. In this perspective, it is believed that being away from the motherland and living in a new culture give authors the chance to criticize the cultural problems of the country of origin, which led to the displacement of the people; the problems that have been carried out to the exile as well with the people. In fact, the focus of the second approach is on the challenges that exist within the culture itself and self-critical reflection is the only answer to the problems. This critical perspective is presented in two ways; either by casting light on the existing problems, examining the reasons behind their emergence and eventually finding solutions for them or by criticizing the current problems by simply verifying the stereotypes that are associated with the South Asians in the dominant discourse. These critics and authors seek to hold a mirror in front of the South Asian culture and tradition and by doing so, portray the pitfalls within the culture, especially in regard to women, as this encounter can act as a self-critical reflection. As Lata Mani argues:

> Questions of tradition and modernity, since the nineteenth century, have debated on the literal and figurative bodies of women. It thus comes as no surprise that the burden of negotiating the new world is borne disproportionately by women, whose behavior and desires, real or imagined, become … the litmus test for the South Asian community's anxieties or sense of wellbeing. For instance, the fear of dating that consumes many South Asian families is primarily a fear of women

dating … it is women who are called on to preserve the ways of the old country

(35).

One of the main challenges of the South Asian culture is the marginalized role of women in literature as well as society and one way that female authors have found to overcome this problem is through opening a new paradigm for representing, negotiating and arguing the positions, concerns and challenges of the south Asian women and have found their place alongside male authors within South Asian Diasporic literature as the main platform to voice their opinions and concerns. Yasmin Hussain, in *Writing Diaspora: South Asian Women, Culture, and Ethnicity,* explores the literature of South Asian women who moved back to India and explores the culture and the tradition of their homeland versus their country of settlement and offers new perspectives on seeing South Asian women. She argues that "South Asian women have reevaluated the concept of South Asian womanhood within the majority and minority cultures as they give voice to their resistance to oppression. The creative input from these women documents this struggle" (Hussain, 1).

Along with race, gender analysis is also another significant theme portrayed in the struggle of the South Asian women against the traditional perspectives, claiming that within traditional societies, gender roles tended to be fixed and communally scrutinized. Some authors, including Bharati Mukherjee and Bapsi Sidhwa, in specific, have focused on the cross-cultural tensions that arise when crossing national borders. Though, this tension is more complicated for women because they must negotiate their ideas broader and deeper to face homogenized perspective toward them (68).

3.6.3 Multi-stream Perspective

In this approach, the critics transcend East-West confrontation and consider Diaspora as a multi-stream process that is not necessarily good or bad, and instead consider it as a dynamic process in which ideas are not in conflict with each other but they embody the intertwined roots of a tree and none of them is superior. Theorists of this approach believe that the time has come to leave the dichotomy of the East and the West, Oriental and Occidental and go beyond them. That is to say these dichotomies summarize a complex set of factors and put them against other. Therefore, summarizing them only based on the power relationship between them, acts as just a simplified narrowed version. Instead, this approach is based on the perception that diaspora is a set of complex factors that is a dynamic process and there is no clear-cut distinction within that. Along with that, authors have also tried to add new definitions to this critical realm. Therefore, a modern understanding of diaspora is encouraged because it goes beyond the illustration of the stereotypes and it serves to break and shift a paradigm that has wrongly framed diaspora as so. In that respect, South Asian American identity should be examined in term of a complex group of social relations and the context that holds and forms them. In this new perspective, South Asian diasporic literature is defined as following:

> South Asian diasporic literature is a diverse and diversely marketed field, and
> both the literature and the discourses surrounding it exist in a determined but
> flexible field of reception. That is to say, while the interpretive discourse around
> South Asian diasporic literature often rehearses similar concepts and tussles
> with the quandaries of self-reflexive critique, the literature itself projects varied
> meanings in the interpretive community (Tamara Bhalla, 134).

Therefore, it is safe to say that South Asian American authors are starting to resist the feeling of either/or between these two fixed worlds and they are beginning to create new spaces for themselves (Amita Shah,96) and through the literary works, they seek to create new possibilities for developing the South Asian American identity beyond the ethnical associations or labels. They also initiated to consider ethnicity as a dynamic process as well. According to Salam, ethnicity is not a static state of being; rather "ethnicity and identity and how they affect life choices are part of an iterative process in which structural contexts and life circumstances influence assimilation outcomes" (8).

3.7 Conclusion

The characteristics of the postcolonial literature and its connection with the Diasporic literature were reviewed in the previous chapter and the focus of this chapter was on the evaluation of the South Asian American Diasporic literature. It was mentioned that South Asian immigrants were considered as a significant ethnical minority in the United States who not only contributed to the economic development of the country, but also played a significant role in the development of American literature. South Asian American authors have contributed to the representation of the South Asian community in the United States and provided a deeper understanding of their diasporic experiences and the challenges and struggles they have faced both in terms of how they are perceived in the host country as well as how they perceive themselves. It was argued that South Asian Americans have been described as the model minority by the dominant voice in the United States to reinforce the myth of American meritocracy. This ideology is identified with the predominance of the white cultural discourse over the embodied other proposed by the Oriental perspective of Edward Said who suggested

that the West has succeeded to form an image of the East that is inferior, backward and irrational and in return presented the West as superior, rational and progressive. The existing stereotypes, especially about South Asian women that originated from the oriental perspective were also discussed and the refusal of the western conception to accept any changes in its perception of the South Asian women was addressed. It was specified that this conception gives the chance to south Asian women to choose between two roles including traditional(backward, oppressive, sexist, uneducated, passive and docile) and progressive (enlightened, egalitarian, active, vocal, and Westernized). Finally three critical approaches toward South Asian diasporic literature were expressed. By focusing on the Oriental perspective of the West toward East, the first approach underlines the employed hegemony by the United States to portray itself as superior and the minority ethnic groups, here South Asians, as inferior. In the second approach, critics and writers challenge their intellectual and cultural heritage and evaluate them thoroughly. These critics and authors seek to hold a mirror in front of the South Asian culture and tradition and by doing so, portray the pitfalls within the culture, especially in regard to women. In the third approach, critics transcend the East/West confrontation and consider diaspora as a multi-stream process that is not necessarily good or bad, and instead is a dynamic process in which ideas are not in conflict with each other and instead they resemble the intertwined roots of a tree in which none of them is superior.

Chapter Four

Methodology and Analysis

4.1 Introduction

      Some critics are still skeptical about the employment of the digital tools for literary purposes claiming that due to the nature of the normal language it can be the subject of analysis for digital tools, but digital methods cannot be employed for literary text because of its unique, innovative and implicit language. Gibbs addresses this concern by saying that the so-called literary language includes "basic schemes by which people conceptualize their experience and the external world" (*The Poetics of Mind: Figurative Thought, Language, and Understanding*, 1). He also adds that literary values are built on various conditions that are previously formed by a given community at a certain time and therefore, corpus analysis can indicate the surrounding of certain linguistic items as employed by authors and therefore enables researcher to provide possible interpretations that could be verified through the text (2).

      The employment of corpus linguistic for literary purposes not only facilitates the identification of the main features of the genre as well as literary style, but also helps to identify

the language pattern of the text by providing valuable clues that are not reachable through traditional methods. This chapter offers the reasons for choosing South Asian American Diasporic Literature as the main corpus of this study and set forth the means that are employed to convert data and engender quantitative data for qualitative examination of the text. In the first section, literary and linguistic criteria for the selection of the subject is presented, followed by a detailed explanation of the design of the corpus. In the next section, the analytical tools and software that were employed in this research are also introduced and data verification and data examination are presented afterward. Then, keywords, collocation, concordance and other lexical bundles that are examined are discussed. Finally, the result of the examination is offered.

4.2 Definition of the Corpus

Corpus is a collection of written texts compiled in electronic format that can be studied to identify the existing trend in the text, emerging formats and to derive a group of invisible rules by which a language is governed *(The Digital Humanities: A Primer for Students and Scholars,* 2). Based on the research questions and the objectives of the study, the scale of the collected texts can vary to be the representative of the specified texts, genre or model to investigate (38).

Douglas Biber et al. address the main feature of the corpus linguistic analysis as analyzing the actual patterns of use in natural texts, utilizing large organized collection of texts as the basis of analysis, making extensive use of computers using both automatic and interactive techniques for analysis and benefiting from qualitative and analytical techniques at the same time (*Corpus Linguistics: Investigating Language Structure and Use*, 4). Based on the main objectives of the research, different methods can be employed to compile a corpus that range

from transcribing data manually to using programming language to mark or tag data for further and more accurate evaluations.

It is important to distinguish between the general corpora and the private corpora. General Corpus includes a large sample of words drawn from collection of texts from different populations. Although it cannot be said that it is the representative of a whole, but it includes as many words as possible within the general use of language, also referred as Reference Corpus employed as a touchstone to compare other specialized corpora with (164-166). General corpora are intended to represent the general usage of language such as British National Corpus (BNC) that includes 100 million words or the Corpus of Contemporary American English that includes 425 million words. Private corpus or specialized corpus represents language usage in specific domains or genres, such as TIME Magazine Corpus, Carnegie Mellon University Recipes or Shakespeare Corpus that are either open source or are available to the audience of specific university, institute or platform. Specialized corpus is the collection of a certain group of texts such as newspaper editorials, academic articles on a certain subject or history books that is designed for a specific purpose (164).

4.2.1 Corpus based vs corpus-driven

It is important to note that researchers distinguish two perspectives toward corpus research including corpus based and corpus driven. As Tognini-Bonelli argues corpus based is an inventory of language data and it is an approach to examine the corpus to answer the raised questions or to test the hypothesis, while the corpus driven analysis is a source of information from which the researcher can retrieve, extract and examine the data to observe the linguistic patterns. In the later approach, corpus is approached with no previous assumptions, while in the

earlier researchers approach the corpus to evaluate pre-existing hypotheses (Corpus Linguistics at Work, 17). According to Chomsky, the difference between two approaches is mainly theoretical. In corpus-based approach, the primacy is given to subjective introspection over objective, empirical data, while the corpus driven approach is the opposite (*Syntactic Structures,*123). This study benefits from an exploratory and analytical approach and the corpus of this research is approached from an objective standpoint without the intervention of the researcher and the of results emerge directly from the employed tools rather than originating from the predetermined ideas of the linguistic patterns.

4.3 The Design of the Corpus

Critics of corpus linguistics have determined certain criteria to consider a corpus accurate and credible which are: authenticity, representativeness, sampling, finite size, machine-readable form as well as having standard reference ("Corpus Linguistics", 1).

Representativeness and authenticity go hand in hand. For a corpus to be authentic to represent a certain population, texts, genre or era it should include a broad group of texts, authors or data to be able to extend the result to the whole group or population. The first consideration for designing the corpus is to determine the right size of texts to be included in the corpus so that the compiled corpus represent the features and characteristics that the researcher has addressed them in research questions. Although the evaluation of the language patterns requires a larger corpus to increase the validity of the analytical results, the restriction on the size of corpus is the result of practical consideration, rather than theoretical issues. The type of analysis also determines the size of the desirable data as well ("Corpora and corpus linguistics", 40).

After specifying the priorities of the research, the criteria for the selection of the texts should be identified. This process in called Data Assignment that is the process for mapping out the type of the data or the text to be covered or collected. For this study, different linguistic and literary factors were influential in narrowing down the idea on which novels to include in the corpus that are referred in the following.

4.3.1 Literary criteria in the selection of the topic

The detailed observation of the diasporic literature shown that the portrayal of an uprooting journey is at the core of all diasporic literature, however the skeptical points of view of the critics have avoided the establishment of an inclusive category that could encompass all the diasporic literature written by authors of different backgrounds. It has also been observed that different factors including forced or voluntary movements, political, cultural and social motifs of the journey as well as the historical backgrounds of the home and host countries, have all given to this prevailing perception that the diasporic journey of each minority group is different from the others; therefore, the literature of each minority group is unique and could not be extended to others to see through their similarities.

It has also been noticed that there is no clear-cut distinction between the diasporic literature of each minority; therefore, invisible boundaries could be set according to the approach of critics and the group to which the writing belong may differ accordingly. For instance, a critic may associate a diasporic novel to a certain geographical setting, while at the same time it could be linked to another group based on the existing similarities in ethnicity, region, culture, language and so on. There are different classifications for South Asian literature, as well. South Asians literature can be evaluated based on its constituent countries individually, or it can be

considered to be a part of Asian Diaspora or it can be associated to the diasporic literature of the country in which it is presented such as South Asian American, South Asian British or South Asian Canadian literature.

The breadth of the classification options, lack of clear-cut distinctions, the resistance toward classifying diasporic literature as a whole, coupled with the unresolved challenges and complexities within this literary arena, have avoided the extension of this idea that Diaspora literature can also be considered as an inclusive genre that is formed by its different subgroups.

Therefore, South Asian American literature was considered as the best example to define diasporic language and its constituting linguistic components. Furthermore, the significant contributions of the South Asians authors to the advancement and popularity of the American Diaspora Literature has been another factor for choosing this ethnic group to be included in the corpus. And finally, since South Asian countries were affected by the colonial process directly, they were among the first nations to address and criticize the impacts of it upon their culture, and therefore their literature can offer a more detailed and accurate picture of the current collective diasporic language.

Different significant factors were taken into consideration in the selection of novels to be included in this corpus. After compiling a complete list of the South Asian authors who have contributed to the diasporic literature of the United States, the literary works that were written in any other languages rather than English were omitted. Also, because the focus of the research is on the original portrayal of the diasporic experience by the authors, those works that were in translation were dismissed. The list of the works then was divided into three groups of novels, memoirs and short stories, but due to the fact that memoir revolves around subjective individuals

and does not necessarily represent diasporic literature it was excluded. There was not also enough diversity in respect to short story, therefore novel was chosen.

To have the most authentic contemporary version of the diasporic literature that portray the current diaspora, a time span between 1980 to 2013 was chosen. Also, given the fact that the female and male authors' authorship style differs significantly, therefore female authors was selected trusting that the works of the female author present an enriched version of the diasporic language portraying their struggle to find a voice for themselves.

Also, given the fact that many of the diasporic novels are infact based on the imagined diaspora, it has been tried to avoid the inclusion of them within the corpus of this study, there was an attempt to choose novels that were written as a result of the real diasporic experience of either the first or the second generations of diasporic communities.

South Asian Diaspora literature includes the literary works of those authors who have moved or descended from India and Indian subcontinent including Pakistan, Bangladesh, Sri Lanka and Nepal. South Asian literature was considered as Indian literature because of the amount of the works and authors that have contributed to that and although it has changed over the time, but Indian authors or authors of Indian descent play a major role not only in the diasporic literature of the United States, but also to the English literature in general. Aside from Bangladesh that did not have any noticeable novel within the designed framework, twenty-five novels were chosen to be included in the corpus as the representative of South Asian literature.

Chart 4.1 Time span of the novels of the corpus
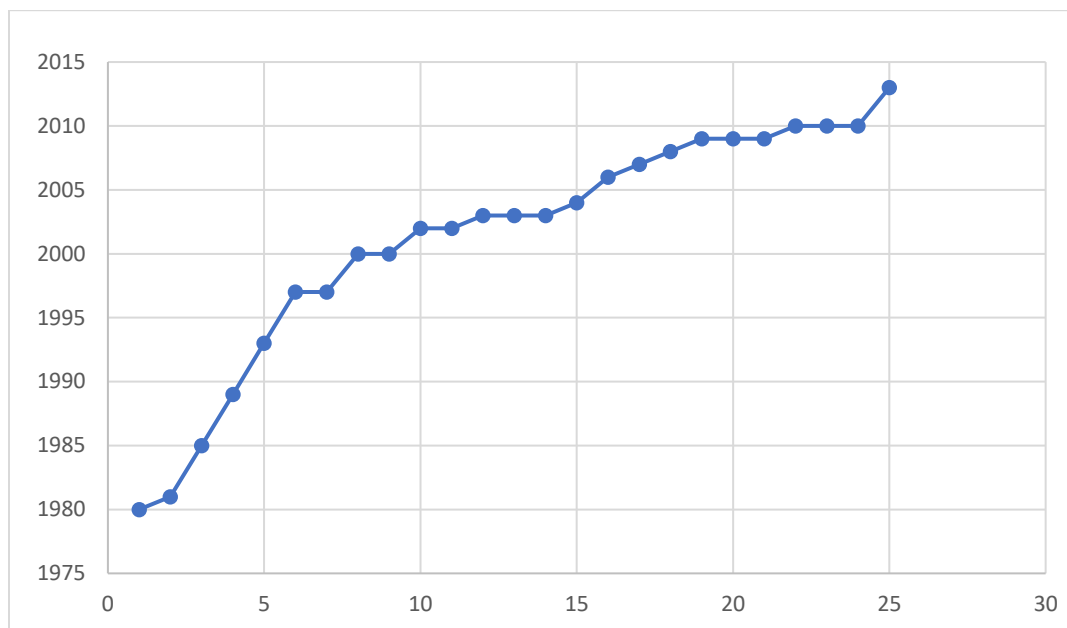
The corpus collection process includes the procedures to turn the physical content into electronic format. For this process, the researcher benefited from the rich libraries of Carleton, Ottawa and Toronto universities. Once the novels were collected, the next step was to transfer them into electronic format by scanning each novel thoroughly through OCR. The process of transforming the novels into electronic format was followed by a two-stage evaluation of the data to check the spelling, capitalization and punctuations both manually and through the available spell-checking software.

Prior to saving the result into electronic format, another step referred as Normalization of the corpus was carried out. The main objective of this process is to modify the measured value of different size and form an average scale. After that the result was saved in txt format as it is the common format that most corpus tools work with.

4.3.2 Text Encoding

As Oakes puts is, corpus annotation or tagging is the process of adding explicit

information to machine readable text (249). The added information is mostly concerned with part

of speech, lemmas or tenses. Text encoding is fundamental in any experimentation of the digital

tools. Sussan Hockery addresses electronic text as "an electronic representation of any textual

material which is an object of study for literary, linguistic, or related purposes." (1) and in this

process, any information that facilitates the performance of the computer program on the given

text is essential *(Electronic Texts in the Humanities*, 1). In this regard, text encoding provides the

tool with the required information to carry out a precise measurement of the text. Encoding of

the texts is mainly carried through XML-based mark-up languages, such as HTML that specifies

a set of rules for encoding texts. The central step in encoding contains tagging, the framework

that initiates and ends with brackets definite elements as attributes are ascribed. Following is an

example of an encoded page:

<!DOCTYPE html>

<html>

<body>

<p>The first paragraph.</p>

<p>The second paragraph.</p>

<p>The third paragraph.</p>

</body>

</html>

4.3.3 Standardized Type Token Ratio

The lexical density of the corpus could be determined by dividing the number of tokens (words without repetition) by the type (unique words), multiplied by 100. Considering that the volume of the texts that form the corpus is one of the most influential factors in determining this ratio, therefore the result could vary very widely in accordance with the length of the text - or corpus of texts -which is being studied. For instance, a text of 1000 words can have a Type-token ration of 40%, while another texts that is formed of 40,000 words, this ratio can be 2%. Therefore the employment of the conventional TTR is beneficial in the analysis of equal size texts within the corpus. As Paul Baker suggests employing Standard Type Token Ratio is more favorable in the evaluation of the corpus of texts of different size(50). In addition to the difference in the size of the texts that form the corpus, the overall length of the corpus is also another factor for using Standard Type-Token Ratio. Baker in *Using Corpora in Discourse Analysis* pointed out to this problem as follows:

> A corpus or file with a low type/token ratio will contain a great deal of repetition of the same words occurring again and again, whereas a high type/token ratio suggests that a more diverse form of language is being employed. … However, as the size of a corpus grows, the type/token ratio will almost always shrink, because high frequency grammatical words like the and to tend to be repeated no matter what the size of the corpus is. Because of this, large corpora almost always have very low type/token ratios and comparisons between them become difficult (52).

To overcome this problem the tool calculates the standardized type/token ratio based on the type/token ratio of the specified number of tokens, which is usually 1,000 and then

calculated afresh for the next 1000, and so on to the end of the text or corpus. Therefore the result is based on the consecutive 1000-word chunks of the corpus. This ratio address the linguistic diversity of the corpus. The rate of the Standardized Type-token is almost always higher than the normal TTR, because of the fact that the focus of the evaluation at each stage is on a smaller amount of text. Low ratio tends to signify that either a small range of subjects are being discussed in the corpus, or in some cases it indicates that the language that is used is rather simplistic.

The lexical diversity of the compiled corpus for this study was evaluated by Wordsmith tool. The corpus of this study includes 2123735 , out of which 60347 are unique which constitutes 2.85% of the content. However the standardized TTR based on the consecutive 1000-word chunks indicate 46.26%. Following table shows the overall information about the corpus and its statistical features:

Table 4.1: Statistical features of the corpus

text file	file size	tokens (running words) in text	types (distinct words)	type/token ratio (TTR)	standardised TTR	STTR basis
Overall	11770808	2123735	60347	2.85	46.26	1000

However, it should be noted that this ratio only gives an indication of the lexical diversity and complexity of the text and along with the above statistics, further evaluations are needed.

4.3.4 Tagging of the Corpus

A preliminary step prior to the practical evaluation of the corpus of this study is to tag the corpus to highlight the linguistic feature of the data. This step adds information about the grammatical category to which every word belongs. The importance of applying Part of Speech tags or what is refereed as grammatical tagging is because authors unintentionally tend to adopt similar syntactic patterns in their writing frequently, therefore the assignment of tags not only helps to map the text in general, but also facilitates the evaluation of the style and the tone of the text. According to Baker, tagging tool designate word categories based on the existing grammatical and morphological rules, patterns, probabilities or the mixture of all the three (*Sociolinguistics and Corpus Linguistics,* 16).

The process of tagging of the Part of Speech (POS) is carried out by TagAnt software. TagAnt is a tool to annotate text with part-of-speech information and is built on TreeTagger that was developed by the Institute of the Computational Linguistics of the University of Stuttgart. The following example illustrates the tagging of the corpus. Tag sets can be found in Appendix i .

Table 4.2: POS tagging of the corpus

```
Zareen_NP Ginwalla_NP hurried_VVD into_IN the_DT hall_NN when_WRB the_DT
bell_NN rang_VVD ,_, waved_VVD the_DT cook_NN who_WP had_VHD popped_VVN
our_PP$ back_NN into_IN the_DT kitchen_NN ,_, and_CC opened_VVD the_DT
portals_NNS of_IN their_PP$ home_NN to_TO her_PP$ husband_NN ._SENT Zareen_NP
never_RB thought_VVD of_IN the_DT entrance_NN as_IN a_DT mere_JJ ingress_NN
._SENT The_DT ancient_JJ door_NN ,_, grooved_VVN by_IN the_DT cen-_NP
turies_NNS and_CC touched_VVN by_IN vestiges_NNS of_IN faded_VVN dyes_NNS ,_,
```

4.4 Corpus Techniques

4.4.1 Frequency List

The first step in the analysis of the corpus is to transform a set of texts into a complete word list. Word list program examines the set of the texts under investigation and reduce the repeated tokens to types. As Scott and Tribbes put it "each instance (token) of the word the is counted but the completed list displays the only once as a *type*, usually together with its frequency (the number of tokens found)." Frequency list or the lexical frequency list refers to the percentage of repetitive words on the corpus. This list includes both the given types as well as the count of how often each has occurred (*Corpus Linguistic Approach to Literary Language and Characterization*, 45).

The result of the keyword analysis could be presented in either frequency order or alphabetical order. Frequency list of the words provides an overall illustration of the corpus and identifies the terms that are common or uncommon in the provided context. According to John Sinclair wordlists are valuable point of departure to familiarize oneself with the corpus and evaluate which words worth further examination (171). According to Scott, automatic comparison of the lexical words with the frequency counts of the reference corpus provides information about the unusual frequency of the words in a given text (236) and therefore it is an invaluable methodology for the examination of the literary text that this research also benefits from. Therefore, the analysis of the frequency list enables the researcher to identify the most recurring words in the given corpus and evaluate them with another corpus to reach an impartial and objective conclusion. According to the classification that is offered by the Longman Grammar of Spoken and Written English words can be grouped in three major classes that include lexical, functional and free moving words. Lexical and function words are present mostly

in academic registers, while the third group is free moving word-like elements that tend to convey emotional and interactional messages mostly in spoken texts (Biber, et.al, 56).

Upon measuring frequent words, one can distinguish between grammatical words (such as *of, the, in, and, a, to*) and lexical words (such as *nice, him, go*). Based on the purpose and the size of the corpus, grammatical words could be ignored by putting them in the "stop list" of the analysis. This step unclutters the wordlist of the issues that are not the main concerns of the particular study (Sinclair, 102). According to Niladri Dash "Lexical words convey meaning in texts, and therefore, provides aboutness information found in context" ("Context and Contextual Word Meaning", 1).

4.4.2 Concordance

Although word frequency provides an objective method of measuring the importance of specific concepts in the corpus, but the representation of the ideas is not usually carried by single words and it is helpful to take the frequency of multi-word units into consideration, what Baker address as clusters (Baker, 2010). This process brings information about the context of the analyzed terms also addressed as the generation of the concordance lines. According to Baker, frequency list is helpful in determining the focus of the text, but it is the concordance analysis that provides the chance for researcher to do a close examination of the text. Concordance, as he puts it "is the list of all the occurrence of a particular search term in a corpus, presented within the context that they occur; usually a few words to the left and right of the search term (*Using Corpora in Discourse Analysis*, 71). Concordance is also referred as key word in context (KWIC). Here, key word simply refers to search term and therefore is distinguished from keyword. Frequency list and concordance are among the most frequent evaluation within corpus

linguistic studies. Concordance lines brings in many examples of the term under discussion. These lines could be ordered alphabetically or based on the context words to highlight structures that convey meaning in the text.

Concordance brings together as many examples of the chosen term as possible. In this stage the words that are taken from the corpus are put back into their context to have a perspective of the word in its surrounding words to uncover their patterns. Node or the word under question appears in the middle of line within its context. It is possible to increase or decrease the number of the words that appear before or after the node. The concordance lines can be sorted alphabetically or by frequency.

4.4.3 Collocation

According to Sinclair, collocation is "the occurrence of two or more words within a short space of each other in a text" (170). Based on this definition, a word collocates with another word if it happens in a certain span from the node, which is the word under question (Scott & Tribble, 2006). Paul Baker in *Using Corpora for Discourse Analysis* puts collocations as:

> All words co-occur with each other to some degree. However, when a word regularly appears near another word, and the relation is statistically significant in some way, then such co-occurrence referred to as collocations and the phenomena of certain words frequently occurring next to or near each other is *Collocation*. As Firth famously wrote: You shall know a lot about a word from the company it keeps". Collocation is therefore a way of understanding meaning and associations between words which are otherwise difficult to ascertain from a small-scale analysis of a single text (95-96).

The analysis of the collocation of the chosen word not only provides valuable information about the pattern of the text, but also highlights the style of the author and his outlook about his or her choices. Collocation demonstrate the relation between terms by evaluating the co-occurrence of specific words together which can unveil details about the connection between them. Although there is no specific indication of the size of the span, but most of studies have focused on 2 to 5 terms on either side of the node, but as Sinclair addressed real collocates cannot go beyond four words from each side of the node.

4.5 The employed Software

Technological progress in storing and processing data have given the chance to compile millions of words, while at the same time it has considerably increased the expectations that are set for corpus linguistics because of the reliable results that these linguistic tools and software present. One of the primary but a significantly crucial feature of the corpus analysis is the presentation of the keyword or frequency list that measures the words in the given corpus, along with giving the chance to examine particular words in their surrounding environment and context. In this study, multiple software are employed that include TagAnt, Lancsbox, Wordsmith, Antconc as well as Voyant and Ubiqu+Ity for visualization of the corpus.

As its name connotes, TagAnt is a free-ware tool for tagging the corpus with Parts of Speech. Antconc is also a corpus analysis toolkit for concordancing and text analysis. This study benefited from the Concordance, Concordance Plot, N Grams and Keyword list functions of the tool. Another employed tool is LancsBox which is an analysis tool that is provided by the Lancaster University and the KWIC of the analyses are carried out with it. Next tool is Wordsmith 6.0 that is a software for mostly corpus linguistic purposes which provide different

modules that can be exploited for different purposes. Patterns and clusters are the two mostly employed parts in Wordsmith. Voyant tool is an open source tool for visualizing the result of the analysis of the corpus that this research also benefited from and the final tool is Ubiqu+Ity, a Visualizing English Print application developed by the University of Wisconsin, Madison.

4.6 Frequency of the Words

Considering the opportunities that the digital-based methodologies have provided in the recent years, the employment of the corpus analytical tools have been on the rise as well. One of the reasons that have increased the authenticity of corpus-based research is the availability of the large corpora that could be employed as the basis for the evaluation of the authenticity of the research corpus, such as British National Corpus or The Corpus of Contemporary American English. The general corpus, also referred as reference corpus provides the text norm and standard against which one can compare the research corpus with and contains millions of words from different sources. However, to address a special issue in the research for which there is no related corpus for comparison yet is one of the challenges that researchers can face. Although it is possible to employ a larger available corpus to address the issue in question.

Since the reference should be a general representation of the language that is questioned, It is fruitless to employ a million-word general corpus as the basis for the analysis of a special issue. For instance, it is inconvenient to use the general corpus to compare a literary corpus with. Given that most of the available general corpora are compiled for linguistic research, the extent of the literary corpora for literary research is very limited; therefore, a specialized corpus of the relevant texts could be built to have a more authentic and accurate results. As the focus of this dissertation is on the particular issue of the South Asian American Diaspora Literature written by

the female authors, a specialized reference corpus was created that included 100 novels of American literature written during 1980 to 2015. This specialized reference corpus was used as the basis of the comparison for the South Asian American Diaspora corpus that was described earlier.

Frequency analysis is practically the first step of the evaluation of the specialized corpus. As Barlowre marks " the most radical transformation of a text used in linguistic analysis is to, in effect, rip it apart to produce a wordlist." (207) Using Antconc wordlist tool, the frequency wordlists of reference corpus and the compiled corpus were generated. Each term in the research corpus was compared against the reference corpus and the program made the decision on whether any significant difference is observed within the frequencies of the both corpora. The frequency list was generated to include all the values, out of which the first 100 words could be found in appendix ii . It is also possible to search for the negative frequency as well, but the preliminary evaluation on the sample of the corpus showed it does not provide valuable insights into research. It is possible to filter out stop-word prior to evaluation, which includes the most common words in language such as *the, a, an* . However, to have an accurate result especially in regard to phrasal verbs, concordance and collocations they were removed manually after the evaluation was carried out. The purpose of this was to keep the focus of analysis on the structures that convey more meaning in regard to diasporic literature.

Given that the constituent texts of the corpus are of different size, the result of frequencies are presented in relative frequency. The convention is to calculate per 10,000 words for smaller corpora and per 1,000,000 for larger ones. The relative frequencies of words for this research was calculated based on one million tokens. Following include the list of the first 15 frequent words across the corpus:

Figure 4.1: Frequency list of the corpus

Term	Count	Relative	Trend
said	9009	4,267	
like	5958	2,822	
time	4048	1,917	
know	3803	1,801	
just	3443	1,631	
mother	3207	1,519	
way	2910	1,378	
man	2875	1,362	
eyes	2712	1,285	
away	2560	1,213	
come	2534	1,200	
face	2470	1,170	
thought	2447	1,159	
day	2379	1,127	
room	2365	1,120	
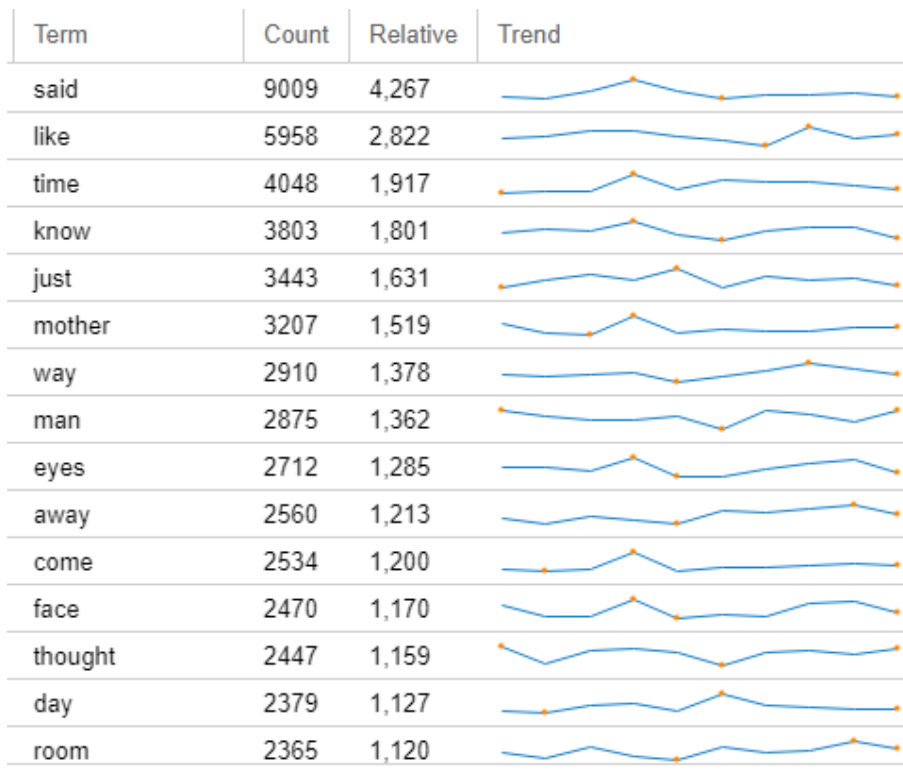
4.6.1 Keyness of the Keywords

As Scott (1999) noted, three types of keywords can be identified within a corpus that include

keywords that provide indications of the aboutness of a particular text, high frequency words that

can address the stylistic features as well as the proper nouns (qtd in "A Corpus-assisted Critical

Discourse Analysis of the Arab Uprisings", 64). AntConc has provided the users with the

possibility of employing both chi-squared and log-likelihood for the keyword statistic test. As it

is argued in *Using Corpora in Discourse Analysis,* the frequency of one wordlist could be

compared with another wordlist to demonstrate which words are statistically more present in the

investigated corpus. In this sense a wordlist provides "*frequency*", while the keyword list gives

a measure of "*saliency*" (Baker, 125).

The test for the keyness value of the keyword was evaluated by the Chi-Squared test. In this study, chi-squared was adopted to generate the keyword and the p value was set on $p < 0.05$ In general the p value of 0.05 indicates a greater confidence in the results that they did not happened due to chance and therefore increase the authenticity of the results.

The keyness could be calculated by carrying out either chi-squad or the log-likelihood test on the frequent words. The comparison of the high frequent words within different corpora could be carried out in two methods; the first type include the comparison of the features of the research corpus to a larger corpus and the second include the comparison of two almost equal sized corpora. To evaluate the distribution of the frequency when comparing a relatively small corpus to a much larger corpus, it is suggested to use chi-squared value to compare the observed distribution with the prediction to validate the results, while Log-likelihood highlights words that are relatively common in general use ("Comparing Corpora using Frequency Profiling", 1-2). The result of the Keyness can be positive indicating unusual frequency of the term in the research corpus, or negative demonstrating unusual infrequency within the research corpus in comparison to the reference corpus. The keyness evaluation is a further analysis to validate the keywords. Following figure shows the keyness evaluation of the keywords:

Rank	Freq	Keyness	Effect	Keyword
1	4055	+ 1843.9	0.0097	time
2	3750	+ 1704.82	0.009	mother
3	2714	+ 1232.88	0.0065	eyes
4	2651	+ 1204.21	0.0063	father
5	2414	+ 1096.36	0.0058	day
6	2361	+ 1072.24	0.0057	house
7	2219	+ 1007.65	0.0053	life
8	2160	+ 980.81	0.0052	head
9	2179	+ 989.45	0.0052	looked
10	2116	+ 960.8	0.0051	hand
11	2106	+ 956.25	0.005	told
12	1934	+ 878.04	0.0046	home
13	1925	+ 873.95	0.0046	tell

4.6.2 The Analysis of the Keywords

The next step after validating the keywords is to categorize the keywords in relation to each other and in connection with the corpus. The distinguished keywords of this research are categorized into three main groups of nouns, verbs and adjectives and each group is divided into further conceptual sub-group to be examined in detail that include:

Table 4.3 The most frequent words across the corpus

Noun
Time , mother, way, eyes, father, face, day ,Room, house, life, head hand ,something, home, look, people, family, woman, love, Padma ,night, things, door, voice, hands, feroza, men, hair, world, child, parents, name, bed, son, India, girl, morning, everything, mind ,car year, course, daughter, school, America, table, light, ellie, moment, friends
Verb
was, is, said, were, are, know, go, see, come, thought, made, looked, say, get, told, think, felt, asked, came, tell, want, knew, going, make, take, left, went, saw, let, wanted, took, turned, got, put, heard, found, says, seemed, called, feel, married, gone, give, stood, talk, began, call, find, help ask
Adjective
More, other, first, few, next, last, same, small, many, young, white, frank, dark, Indian, black, better, red, big, full, high, great, whole, blue, second, happy, bad, cold

The initial overview of the nouns points out some relations to diasporic features of the corpus. The very idea that the lexical term "*time*" is one of the five most frequent word within this diaspora in itself highlights the importance of the concept of time in diaspora. Those in diaspora develop not a physical, but a sentimental bond with the concept of time recalling the old time and days. Other frequent terms that refer to time are *night, days, morning, yea* and *moment*.

The concept of family and its constituting members is also another significant features of the diaspora which is also seen in the frequency list of the noun. The quest for keeping family together, which is a big part of the representation of the home is a central part of the diasporic life and the evaluation of the concordance lines of these words can show if these words are employed with the same viewpoint or not. The family related frequent words are *mother, father, woman, men, husband, wife, son, daughter,* and *girls.* Within diaspora, the role of woman as a mother comes at the core to keep family together. As the chart shows, even the role of a woman as a wife comes second to motherhood. It is also the same about men, the being a husband is ranked as secondary in this division. Although the visualization of the corpus indicates that father and mother start almost a same level, but at the end the distance between motherhood and other roles remains equal.

Chart 4.2 Distribution of the family roles
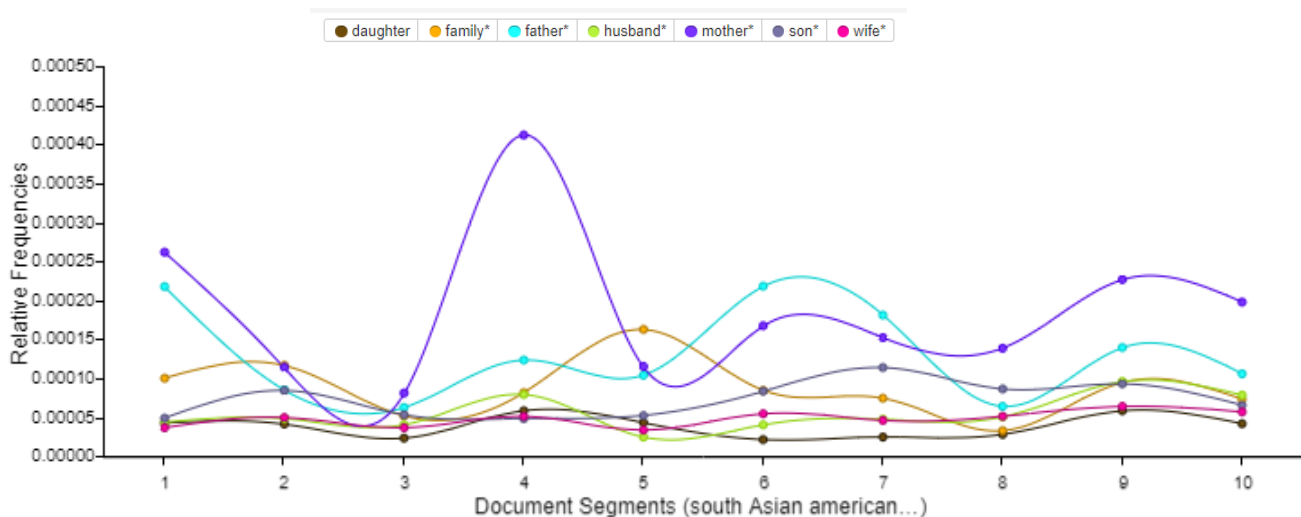
Lexical terms that are related to appearance including *eyes, face, head, hands, voice, hair* could be grouped together for find out if they are related to diasporic literature. Next group included

the lexical terms that referred to either a location or object and include *room, house, door, bedroom, India* and *America*. Words including *way, life, home, people, love, world* and *mind* are considered as abstract concepts. The visualization of the frequency of the abstract nouns represents interesting facts about the diasporic corpus of this study. *Mind* which can be the representational of the rational thinking comes at the lowest level of this group indicating that it is not functional in this structure, *Way* and *love* start off together, accompanied by *home* in the middle. The hesitation between *way, life* and *love* continues. The final point in which *life* and *home* come close to one another underlines the fact that to have a *life* and to be able to *live,* it is necessary to seek home (that can be the indication of the identity or homeland).

Chart 4.3 Distribution of the abstract concepts across corpus
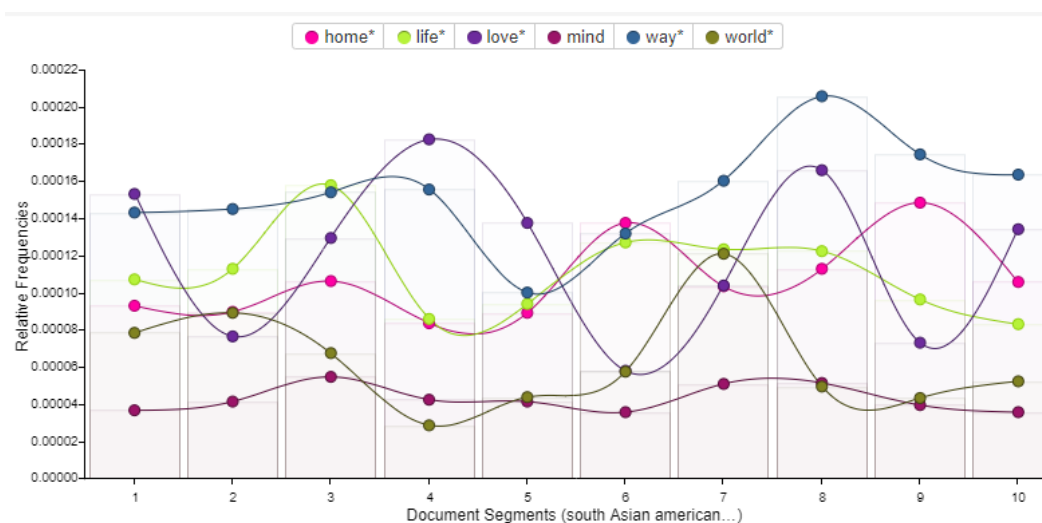
Within the classification of the verbs, the focus was on the words rather than the lemma. lemma is the canonical form of a word and as Francis and Kucera define lemma is a "set of lexical forms having the same stem and belonging to the same major word class, differing only in inflection and/or spelling" (*Frequency Analysis of English Usage: Lexicon and Grammar*, 1).

For instance *saw*, *seen* and *seeing* are considered different words, but all of them are rooted in *see* as their lemma. It is possible to add the lemma list to the software, so that the words with the same base are treated the same. But given the fact that different structures of the words, especially verbs, provide valuable information about the text and its relation to the subject of time in diaspora, words and not lemmas were chosen as the basis of the analysis. However, to provide a clear category of the frequent verbs, at the categorization level the basic form of the verbs is considered.

The verb *said is* the most frequent verb within this corpus, followed by *know*, *come* and *thought.* Living in between two-cultures is a hard experience for those who are in diaspora, but feeling of not belong to the host land and losing the homeland gives a space to them to talk about their experience and share them. This Third Space gives the chance to them to speak and that is why *said* is the most frequent word of the corpus, as diasporic literature is the third space to talk. Other verbs in the same group as *said* are *talk, ask, tell* and *call*. The rest of the frequent verbs include stative verbs such as *see, know, think, feel, look, seem, let, like and want* and only four verbs are active verbs including *go, put, come* and *make.*

Within the adjective category, *more* with relative frequency of 1,496 is the most frequent one. *New* and *old* are two concepts that are common in diasporic literature in reference to home, culture, identity and setting, which are also among the most frequent adjectives of the corpus. Adjectives could be classified in almost three categories; descriptive, comparative and quantitative. Descriptive adjectives include *small, young, dark, black, big, full, happy, old, new, bad, cold* and *good* that explain the state of the being of a person or object. Comparative adjectives include *more*, *same* and *better* and quantitative adjectives contain *firs, next, last* and *few.* Following is the chart of the most common adjectives within the corpus of this study:

Chart 4.4 Frequency of the Adjectives

Relative frequency of Adjectives

The chart shows a descending bar graph with the following adjectives on the x-axis: more, other, first, old, new, good, few, next, last, same, small, young, dark, black, better, big, full, happy, bad, cold. The y-axis ranges from 0 to 1600.
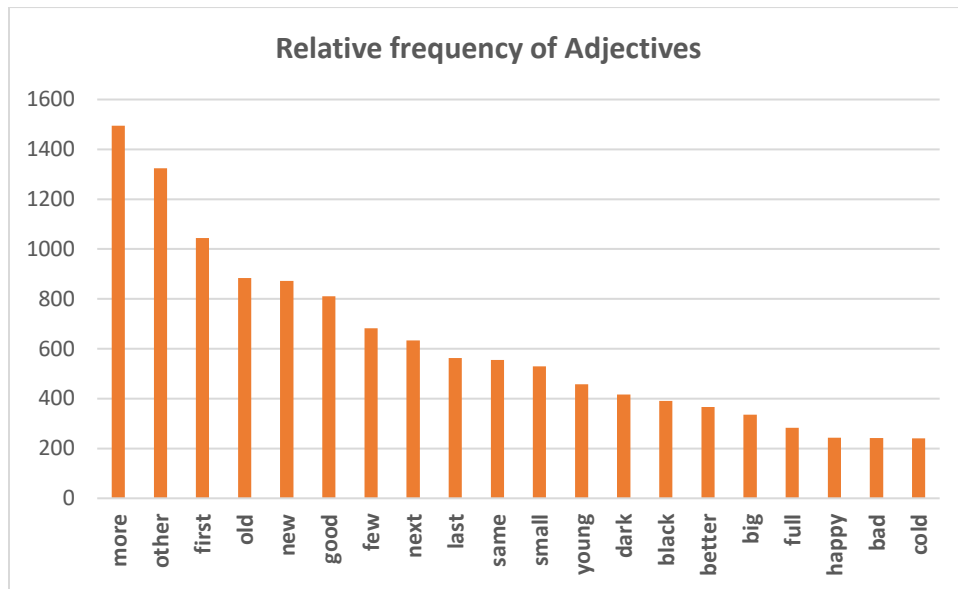
4.7 Concordance Analysis

Given that a general picture of the corpus and its main constituent words is portrayed in frequency analysis, the next step is to put those term back into their context and examine their behavior within their surroundings. According to Baker, concordance analysis is one of the most effective techniques which allows researcher to carry out a close examination of the text and spot the language pattern (71). By providing the context in which the word occur, concordance analysis provides further insight about the text that should be uncovered. In concordance analysis, the main lexical term under investigation is referred as *Node* and the words before and after that provide further information about it.

Figure 4.2 Concordance lines of the *know*

Index	File	Left	Node	Right
1	1. An America	phone anymore! What if it's some-one I don't	know?"	Zareen mimicked her daughter in English. "I
2	1. An America	removed my cardi0an. Feroza pretended she didn't	know	me. "In the car she said: 'Mummy,
3	1. An America	197B! Our Parsec children in Lahore won't	know	how to mix with Parsee kids in
4	1. An America	to control law and order! Didn't he	know	he was inviting martial law Nationalizing even
5	1. An America	to by the fundos," Zareen retorted. "You	know	what he said when they accused him
6	1. An America	of the servants in the kitchen. "I	know	you think my sari-blouses are short, but
7	1. An America	able to act in the play. You	know	how it is -my father won't like
8	1. An America	boys -or any boys. Zareen did not	know	the way the men talked about women.
9	1. An America	his emotion, shouted, "So what Don't you	know	the bastard had drawn up a hit
10	1. An America	earlier. "I'll look after her. Let me	know	when she's coming. I'll go to New
11	1. An America	shriek into the receiver. "Stupid girl. D'you	know	how much your screeches are costing your
12	1. An America	said Cleopatra bathed in asses' milk didn't	know	what they were talking about. The daily
13	1. An America	its natural cycle of six months. "I	know	a remedy," the ayah piped up, transferring
14	1. An America	agreed, "but the only sure cure I	know	is to tic a black thread on
15	1. An America	is turning her quite strange. And you	know	how Bhurro's trial is getting to her."
16	1. An America	of her tragic gaze, she commented, "You	know	how it is. It is not 5ood
17	1. An America	because this is how you talk!" "You	know	I'm always polite to her. Next time
18	1. An America	effort I make." "You think she doesn't	know	how you talk behind her back? She's
19	1. An America	I will do what is right." "You	know	Cyrus loves to see you," Zareen said
20	1. An America	don't care much for old-fashioned thinking; you	know	that." "Yes, yes. You and your Ping-Pong
21	1. An America	she required from Feroza. She must also	know	how eager Feroza was to go, and
22	1. An America	particulars of each. Like most Parsees, who	know	very little about their religion, Feroza had
23	1. An America	in the teeming world she wished to	know	and love as much as she loved

Analysis is carried out by Wordsmith and Antconc tools in two levels; first, the analysis of the concordance lines of the chosen frequent words was carried out and then, the collocation analysis of the terms that the researcher considers to be helpful in unraveling the structure of diasporic literature was completed.

The concordance analysis of *India* and *Indian* reveals interesting features. There is a great tendency to generalize the attributed feature of Indian to all of them, For instance *all Indian parents, all Indian servants.* The only adjectives used prior to Indian even referring to their sales is *aggressive Indian sale* and *stupid Indian habit*. Whenever there is a reference to India the focus is either on returning to India or talking about how things are or done in India.

The close examination of the concordance lines that contained the word *mother* revealed a strong connection between *mother* and *home*. The analysis of the concordance lined showed that *Mother* is closely attached to the concept of *home* and whenever there is a reference to *home, mother* comes afterward. Even the cluster of *"coming back home"* is mostly followed by *" to*

97

mother." Mother is closely associated to home and home is the most frequent location she is

referred to where she is either occupied doing something or her situation or behaviors are

described in home . busy doing something.

Figure 4.3 Concordance lines of mother and home

were strange and strong. At home, my	mother	burned candles so that the house would
had come home gleefully and told her	mother	all about it. Laughter had threatened to
accompanied by his father. At home, his	mother	was weeping because she had not estimated
parchment in his lap. At home, his	mother	is horrified. What type of field trip
he is home for the weekend, his	mother	brings it up again. This time, because

The evaluation of the concordance lines also shows that there are three modes of referring

to mother's situation; her being in the outside world (society) her existence the inside world

(home), as well as the general assumptions about mother. However, the portrayal of *mother* in

the first two types in specific are different; while she is pictured as a passive, gentle and in need

person who should be supported in the outside world, in the inside world she is active and

demanding. In the earlier she is *surviving, widowed, scared, ill, having gentle voice, never been*

to dentist, while in the later she is *asking, urging, wanting*. The analysis of the concordance lines

in which the term mother is mentioned also indicate two types of relationship between the

portrayed mothers and children, in which the first type is dominant. First type showed a

challenging relation between mother and her children in which she is either unable to

communicate with them properly and misunderstand them, or she tries to impose the traditional

perspectives or ideas on them.

Figure 4.4 Concordance line of mother

away from home, a place where my	mother	couldn't always dominate and coerce. A home
he is home for the weekend, his	mother	brings it up again. This time, because
You can stay at home with your	mother.	No schooling for little Zahida, all right?'
myself? My teeth are perfect. My	mother	me to smile with them showing. But
against the rosewood dresser, and when her	mother	caught hold of her wrists, she tore
Sandhya pulled away and started striking her	mother.	First she hit her on the broad
myself I had to do exactly what my	mother	wanted me to do. Maybe that, I thought,

Following also shows some of the described the embedded cultural expectations of mother in the corpus:

Figure 4.5 The image of mother in the first model

Left	Node	Right
"There's one small thIngi" he said. "My	mother wants	to call you Nandini. She doesn't like
asked. "I have no idea why your	mother wants	us to do what she wants us
say anything incriminating? "They're leaving and your	mother wants	you there," she said, and then smiled
accelerated certain expectations, that by now his	mother wants	him settled. The fact that he is

The second type of the pictured relationship with mother is the positive one in which mother tries to keep the family together and sacrifices herself to save family, children or she represents the missing or the longed values of the culture in diaspora.

Figure 4.6The image of mother in the second model

kill. My **mother** was a sniper. **She wanted to spare me the pain** of a

pain of a dowryless bride. **My mother wanted a happy life for me**. I survived

I want to belong. I wanne a mother, a family, Aunties, Cousins. "Where were

sighed, **first my grandmother and now my mother. It was a family thing**, probably embedded

most petty of ways towards Taj. **Her mother saved the family**, in a manner of

are not alone. **You don't have your mother, but, you have your mother's family** to

The final evaluation of the term *mother* was to see what assets are attributed to her and for doing so the term *"mother's"* was evaluated. The most frequent words that collocated with the term mother were mother's *voice, hand, face, touch, lap, presence, sari* and *family* that are mostly the feeling that are contributed to mother.

The next word that has been reviewed is *home.* The evaluation of the concordance lines of the term home showed that it is mostly a state of being, a feeling and a quest for identity and belonging. Some of the examples include:

Figure 4.7 Concordance lines of home

Tears trembled on her eyelashes	Home	is where the heart is, and you
that would not be pacified. Nothing of	home	or motherland pierced his thoughts1 for such
the sentence she had penned. The word	home""	scrawled as if the hand hadn't quite
the original dining room, was now Didi's	home"	away from home, the "Na- maska.r

The word *home* was mostly co-occurred with the lexical term *make* and *build* connotating the necessity of *building a home* which is mostly a sensational one and completely contrasted with the *house* that is mostly referred to the physical nature of it and its geographical location is indicated, that is why it is mostly collocated with *buy*. The difference of *house* and *home* is best expressed by the following concordance line *" to make a house, home."* The evaluation of the concordance lines also showed that the definition of *home* in reference to *women* is constrained by the being with her *husband*.

Figure 4.8 Concordance lines of husband and home

only a year. I would have had a husband, a place to call home. This, Irealize, is
week," Amit said. "There's no privacy at home." To please her husband, Dimple took to wearing
, she knew that her place was with her husband. That her home was where he was. That
York \x96 returning in time to welcome her husband and son home. \x91Do you still think
had called her forth. A daughter's only home is with her husband, that was what was
the kitchen, and opened the portals of their home to her husband. Zareen never thought of the

4.8 Collocation Analysis

Collocation as Stubbs puts it is the frequent co-occurrence of terms in natural language (29). The evaluation of the collocations yields interesting facts about the existing relation between the words and facilitates the execution of a precise analysis on the given linguistic item. Collocate evaluation mostly take plus/minus five terms into account to carry out the collocation analysis. In this section, the collocation of the terms *moment* and *world* are evaluated to see if they provide any further information about the linguistic or literary structure of the diasporic literature, which is the main focus of this study. The collocation analysis was carried out using collocate and clusters/N-Grams functions of AntConc tool. The evaluation of the term *moment* showed that *moment* is mostly located within the cluster of "*for a moment*, *for a moment he*, and *for a moment she*". The close examination of the concordance lines of these clusters pointed out that *for a moment* indicates an astonishment or a moment of wondering, however after that moment of astonishment the male characters (*for a moment he*) are unable to do anything such as *for a moment he was speechless, wondered, gazed, thought, hesitated*, but after the moment of astonishment the female characters (*for a moment she)* are capable of reacting, The examples are; *for a moment she slowed, allowed, wanted, heard, decided, let*.

The next term that was evaluated in this section was the word *world* which mostly collocated with *new, other, difficult* and *different.* Also, *the world of* is the most frequent cluster that is followed by three types of descriptions. World of positive connotations including the *world of books, communication, civilization* and *possibilities*. Next, a world of negative concepts such as *the world of fatal hates, fear, war* and *lies.* The third category includes the world of unique names including *the world of Ami, Kiran, moholla* and *dilli*. The close evaluation of those lines showed that the first type is the *world* that people are interested about and it is explained with verbs such as *live in, return to, lead to* and *dream* that highlight the passion of the people toward that. The concordance lines of the second group indicated the attempt for keeping distance from it and verbs such as *move out of it*, *battle it* and *brought outside it* stress this issue. The last *world* is a world of the past that people *still think of, return to* or *are in contact with.*

4.8.1 Concordance Plot Analysis

Concordance plot presents the searched word within the context in the vertical format and by doing so visualize the times that the word under question is repeated in the corpus. Concordance plot or dispersion highlights the setting of the term and can be employed for a corpus or a set of corpora. In this part, the visualization of the term *but* is carried out through concordance plot function of Antconc. The reason for choosing the term *but* lies within the idea of Homi Bhabha that those he addressed that those in diaspora seek to be like the natives by imitating their ideas, behaviors and dresses and they succeed to be "almost the same but not completely." From grammatical point of view, *but* is a conjunction that can be used to show the contrast within the ideas of the same level. Here is the dispersion of the word *but* across the novels:

Chart 4.5 Dispersion of *but* across corpus


```
                                                        Hits: 619
                                                        Chars: 635453

Plot: 10    FILE: 16.
                                                        Hits: 417
                                                        Chars: 584501

Plot: 11    FILE: 18.
                                                        Hits: 546
                                                        Chars: 538105

Plot: 12    FILE: 19.
                                                        Hits: 652
                                                        Chars: 626152

Plot: 13    FILE: 20.
                                                        Hits: 701
                                                        Chars: 627887

Plot: 14    FILE: 21.
                                                        Hits: 427
                                                        Chars: 479148
```

The visualization of the term *but* represents its dominance over the diasporic corpus showing a never-ending contradiction. The constant presence of it also implies the impossibility of overcoming this contradiction. Here are some of the examples:

Figure 4.9 Concordance lines of But

my life, I can't live without her   but   Acca, she too is consuming me. Listen

That ha didn't question tradition but accepted it just the way he accepted

Frankie'd looked exotic, but   acted familiar. Ham and Larry were harder

will have their own then," I joked but also knew it was true. My grandfather

you and Rabi would not discourage me, but   am I a man who has failed at

no public proof of Rajan's involvement,   but   all his friends knew he was involved,

The evaluation of the collocate of the word *but* also indicates that *he,* and *I* are the two most frequent words that collocate with *but.* This issue, especially in regard to the person himself represent a constant process of questioning oneself, its choices, its identity and its personality. It can also indicate the contradiction that is embedded in the life in-between.

4.9 Conclusion

This chapter included reviews on the definition of corpus, its constituting elements, the design of the specialized corpus for this study, as well as the evaluations that have been carried out on the corpus to answer the research questions. Corpus was defined as a collection of the compiled texts in electronic format that can be studied to identify the existing trend in the text, emerging formats and to derive a group of invisible rules by which a language is governed. It has been discussed that the required corpus could be designed if it does not exist already, but it should have certain criteria to be considered accurate and credible that include authenticity, representativeness, sampling, finite size and having machine-readable format. By taking these factors and other literary considerations into account, the corpus of this study was designed and implemented.

To have the most authentic version of the diasporic literature that portray the contemporary features of diaspora, a time span between 1980 to 2013 was chosen. Once the twenty-six selected novels were transformed into electronic format and the POS tagging and normalization were carried out, the evaluations were performed using corpus linguistic tools.

These assessments included the evaluations of the keywords, concordance lines, collocations, concordance plot and visualization. The evaluations showed the significance of time, family relations, objects, locations and abstract concepts within South Asian American diasporic literature. It was also shown that living in the third space gives the chance to those in diaspora to talk and share their experience, thoughts and ideas of living in-between two cultures and that is why the most frequent verbs include those that are concerned with narration including said, tell, and talk. It was also revealed that there is a close connection between the concept of home, mother and whenever there is a reference to mother, there is a reference to home as well. The word home was mostly co-occurred with the lexical term make and build connotating the necessity of building a home which is mostly a sensational one and completely contrasted with the house that is mostly referred to the physical nature of it and its geographical location is indicated, that is why it is mostly collocated with buy. The presentation of the image of the mother within the corpus included two main portrayal of mother as in the outside world and in the inside world. She was presented as a passive, gentle being in the outside world of society, but within the inside world of home she was pictured as active and demanding. Finally, the evaluation of the distribution of the term "but" indicated the contradiction that is embedded in the life in-between.

Chapter Five

Conclusion

5.1 Summing Up

       This study sets out to investigate the linguistic and literary characteristics of the South Asian American Diasporic literature. The answers to the following questions constitute the main part of this study:

       1. Does the employment of the corpus analytical tools uncover the central framework of South Asian American Diasporic literature?

       2. Is it possible to identify the linguistic and literary structure of the Diasporic literature?

       3. How have the authors adopted a certain terminology to convey their messages and whether the identification of these patterns determine new findings about diaspora ?

       4. Whether the findings of the corpus-driven approach add to the current literary views on the Diasporic literature or dismiss them?

       This chapter encompasses the summary of the crucial factors already mentioned, along with providing the contributions of this reading as well as the findings and suggestions for

further readings. The overriding purpose of this study was to launch a critical investigation of the compiled corpus of Diaspora by employing the corpus analytical tools.

The first chapter included the elaboration on the emergence, development and the prospects of the digital humanities. Given the interdisciplinary nature of digital humanities, it has been discussed that scholars of different backgrounds have tried to provide the inclusive definition of the digital humanities, as well as offering the exclusive definition of that within their field of study. It was pointed out that all the definitions have similar cores including the focus on the digital data rather than the physical text, employing tools for reading and interpreting data, emphasis on the collaborative nature of study and being open to new methodologies and methods. Next, it has been suggested that three eras for digital humanities can be considered both in terms of time and concept. It has been argued that the early prospect of the digital humanities started with the study of the works of Thomas Aquinas through computing and sought to show how the movement from humanities computing to digital humanities took place.

The second phase included the present prospect in which scholars attempt to explain how the digital tools and methodologies could be employed and the final phase includes the future prospects toward humanities that underline where digital humanities will go from here and what kind of possibilities it will provide and what can be expected from it in future. Then, the ways through which literature has benefited from digital humanities have been discussed and the types of research that have been carried out so far in this respect were discussed. Finally, it was mentioned that the rise of distant reading and literature at scale were two significant accomplishments of digital humanities in literature, as they changed the methodology of research, types of the questions that could be asked and enlarged the scope of the context that could be evaluated.

The focus of the second chapter was on the idea of the colonial discourse and the way it led to the rise of the diasporic discourse as a way of resistance. It has been discussed that the ambiguity of the term "postcolonial" resulted in the continuation of the dominance of the colonial discourse in the new form. Also, it was argued that the provided definition has ignored any reference to the local inhabitants prior to the colonial supremacy or the confrontation of the coloniser and the colonized. It was explained that the advent of the postcolonial studies provided the platform for questioning the concept of the West all over again and it has made it possible to evaluate the significance of colonialism and imperialism on the emergence of the West and its perception of itself and the world around it.

Within that framework, the idea of Edward Said, Homi Bhabha and Gayatri Spivak as the "Holy Trinity" of the postcolonial studies were addressed. It has been discussed that Edward Said considers the relations of the East and the West or what he refers to as the Orient and the Occident as a relation of power, domination and different degrees of complex hegemonies, in which the Western knowledge of the East has led to the domination of the West over the East. Rejecting the absolute dominance of one over the other, Homi Bhabha's point of departure from Said's idea was addressed as that he finds the relationship between the colonised and the coloniser more complex and he believes that there is a space from which the colonised can not only communicate his experiences in form of a concept to the coloniser, but also resist domination. And finally, the definition of the diasporic study as an integral part of the postcolonial studies was proposed and it was indicated that unlike the oriental studies that focus on the perception of the West from the East, the focal point of the Diasporic studies is on the experience and the reading of the Orient, along with offering the possibility of examining the idea of self, identity, space and belonging in relation to one another.

The third chapter included the reasons and the significance of the migration of the South Asians to America as well as the contributions of the diasporic literature of the South Asians to the American literature. It has been discussed that South Asians constitute one of the biggest minority groups of the United States and therefore, the model minority stereotypes are attributed to them by the dominant hegemony. It has been defined that the main objective of the minority stereotypes is to portray the included members as a homogenous ethnic group and exclude those that do not fit in. Therefore, within this perspective, people are expected to meet a certain defined social and political agenda that makes it harder to fit in. It was mentioned that the diasporic literature of the South Asians is also expected to represent the stereotyped image of the South Asian American people. By focusing on the diasporic literature of the South Asian female authors, it has been discussed that authors take three approaches toward the addressed stereotypes in their writings. In the first approach, authors focus on the Oriental perspective of the West toward the East and they seek to underline the employed hegemony by the United States to portray itself as the superior and the minority ethnic groups, here South Asians, as the inferior. This approach focuses on the doctrines of Edward Said and other critics to criticize the ways in which South Asians have been employed to strengthen the American discourse and authenticity. The second approach is a self-critical one in which it is believed that living in the third space has given the chance to authors to focus on the challenges that exist within the cultural heritage of them and to criticize them thoroughly. The final approach is a multi-stream perspective and therefore South Asian Americanness is expected to be examined in terms of a complex group of social relations.

The fourth chapter covered the methodological as well as the practical analysis of the corpus. It included the definition of the corpus as well as the differences between the corpus-

based and corpus-driven analysis. The general criteria for designing the corpus have been discussed and the literary as well as linguistic factors that were influential in the selection of the South Asian American Diasporic literature were explained. The following steps of the formation of the corpus were discussed that included the scanning of the texts, transforming them into the electronic format, checking the format and the content and saving the result in the txt format, along with tagging the corpus with Parts of Speech (POS). After providing the definition of the keyword in the context, wordlist, concordance lines and collocation, the analyses that had been carried out on the corpus were explained. The obtained results were introduced and then categorized into three main groups of nouns, verbs and adjectives. Finally, by employing the visualization tools, the results were portrayed.

5.2 Findings

The main objectives of this study can be categorized into three main groups; first, by employing the corpus analysis tools, this research sought to provide a new reading of the diasporic literature, especially within the South Asian American community. Second, this research offered a further method of analysis that is constituted of the linguistic as well as literary evaluations together to provide a comprehensive definition of the diaspora. And finally, the obtained results in this study were compared with the traditional diaspora theories and criticisms to find whether the modern readings add to the traditional readings of literary texts or whether digital studies should be considered as the point of departure from the traditional ones. In this regard, the findings of the research will be presented in three parts covering the linguistic and literary findings as well as the comparison of the results with the current literary theories of Diaspora.
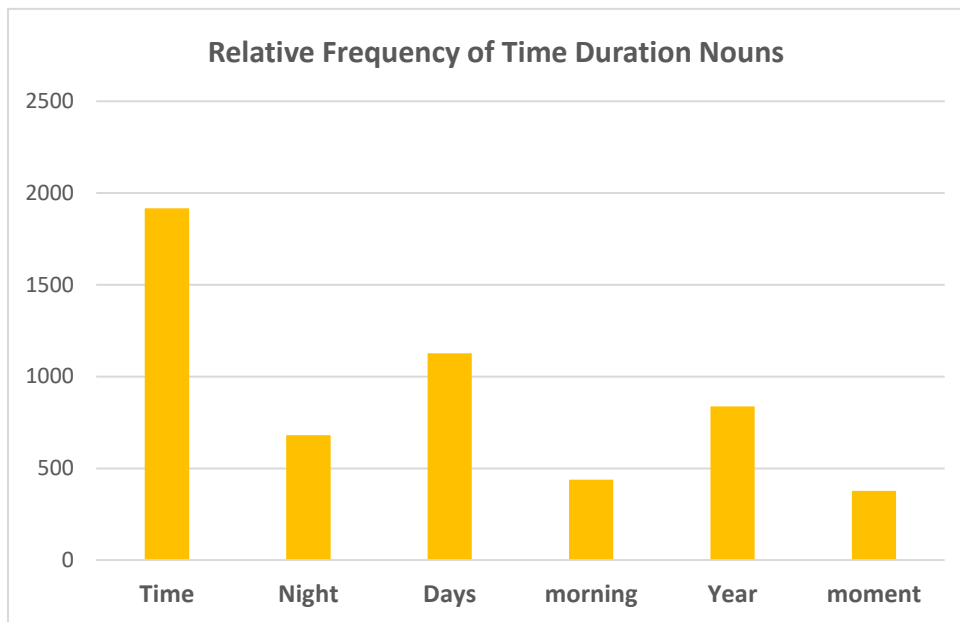
The main objective behind evaluating the Type-token ratio as well as its standardized rate per 1000 words was to identify both the diversity of the words that are employed within the corpus, and evaluate the uniqueness of the text. However the result of this ratio alone does not indicate the uniqueness of the corpus, and this ratio along with other assessments including frequency, keywords, concordance and collocation could demonstrate the exact degree of text uniqeneess. The diversity of the words within the compiled corpus of this study was 2.85 and the STTR was 46.26. The preliminary evaluation could indicate that the low rate of the diversity of the lexical terms can indicate that the significance of the diasporic literature is not due to its diverse lexical range, but it lies within the diasporic meaning it seeks to convey to readers.

To identify the main features of the compiled corpus, a group of analysis have been carried out that include wordlist, keyword in context (KWIC), collocation, concordance and the concordance plot and the results were portrayed by the employment of the visualization tools. By using Antconc wordlist tool, each novel was turned into its forming lexical terms and the repetition rates of them were specified. Upon comparing the frequency list of each novel with the rest, a further list of the frequent lexical word across the corpus was compiled and it was categorized into three groups of nouns, verbs and adjectives. Each group was then divided into subgroups for further analysis. Nouns were divided into five main groups including *time duration, body, family, objects and locations* as well as *abstract concepts*. Verbs were divided in *narrative, stative* as well as *active verbs* and finally, adjectives were categorized into *descriptive, comparative* and *quantitativ*e adjectives.

The first provided category for nouns includes *time.* The significance of the concept of the time within diasporic literature is highlighted by the fact that the term *Time* is the third

frequent word within the corpus of diaspora. The following chart shows other words in this group:

Chart 5.1 Frequency of the time-related words

Relative Frequency of Time Duration Nouns

A bar chart showing relative frequencies for the following words: Time (approximately 1900), Night (approximately 680), Days (approximately 1120), morning (approximately 440), Year (approximately 830), moment (approximately 375). The vertical axis ranges from 0 to 2500 in increments of 500.

The evaluation of collocations of these words as well as the analysis of these concepts within their context indicated a negative connotation of time. After examining the concept of time in the Corpus, It was concluded that three states of time are frequently used within the Diasporic literature: Time as a reference to the homeland and the old country which is mostly a sensational one and recalls the good time of the home country, *Biographical Time* that pictures the experience of the characters' main life events, and the final perception of time within the corpus includes the *Subjective Time* that is the individual's perception of the concept of the time. The subjective time usually come along with the *Descriptive Adjectives* that indicate the

understanding of the individuals of the time. Descriptive adjectives that collocated with the *subjective time* in the corpus include *young, dark, happy, old, new, bad, cold* and *good* .

The concept of family and its constituting members is also another significant feature of diaspora which has also been portrayed in the frequency list of the noun. At the core of the family in diaspora, it is the *mother* that plays the most significant role:

Chart 5.2 Frequency of the family-related words
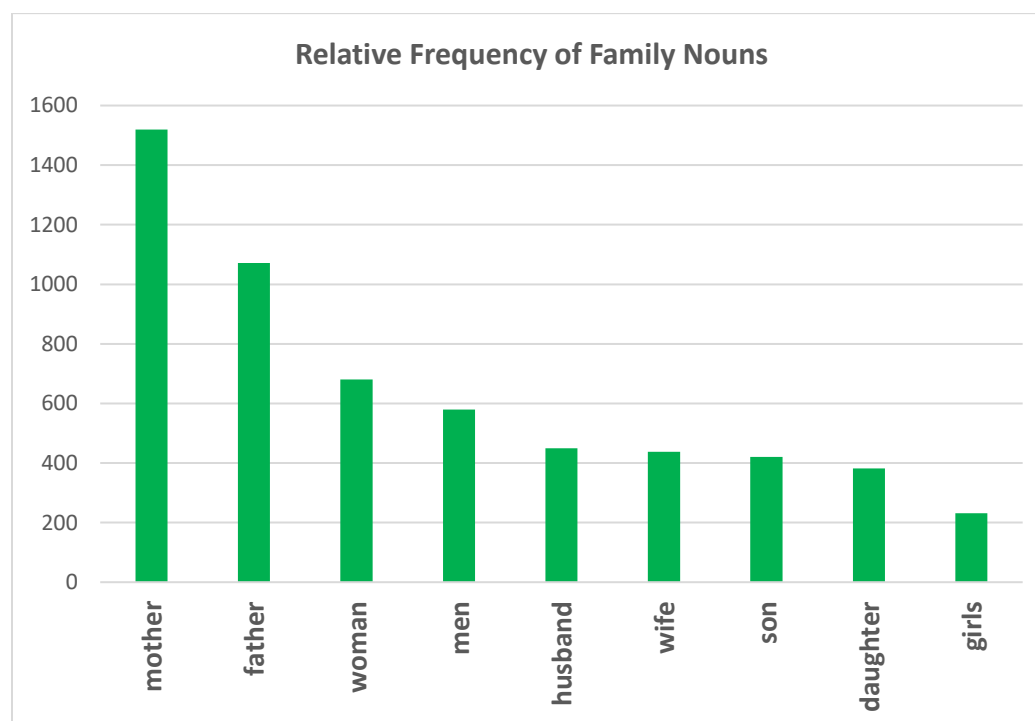
As indicated in the chart, the repetition of the concept of mother almost three times more than the concept of *wife,* highlights the im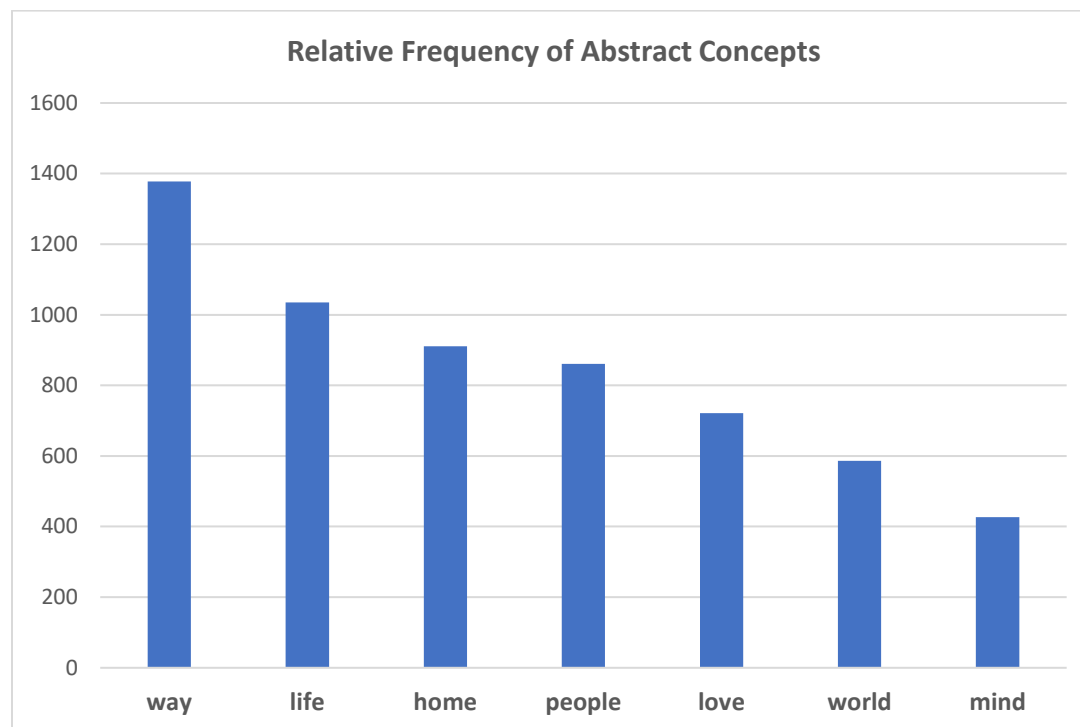portance of the motherhood in the Diasporic literature. In other words, motherhood not only comes before wifehood, but by referring to the emphasis of the texts on it, it can be also claimed that it comes before womanhood as well. The quest to keep

the family together is also a significant element of the representation of the home and the analysis of the concordance lines of the words related to family confirmed that as well. The comparison of the concordance and collocations of the terms *mother* and *father* indicated that mother is mostly linked to home and conveys the sense of home, while father is mostly connected to house.

The evaluations also showed that whenever there is a reference to *home*, there is also a note or statement about *mother* as well. The evaluation of the concept of mother indicated two types of reference to her that included her being in the outside world (society) and her existence in the inside world (home). In the first type mother is pictured as passive, gentle and in need that should be supported, while in the inside world of the home, she is active and demanding. The examinations also showed that the wo types of relationship between mother and children are portrayed within the corpus. The first type portrays a challenging relation between the mother and her children, in which she is either unable to communicate properly with them, misunderstand them or imposes traditional doctrines on them. The second pictured relation is based on the mother's quest to keep the family together and her sacrifice for the sake of children and family and therefore, she portrays the missing values of the homeland within diaspora. The evaluation of the words that collocated with the term *mother* also showed that it is mostly collocated with the words that convey feeling and sensation such as mother's *voice, hand, face, touch, lap, presence, sari* and *family*.

The next category of the nouns included the words that conveyed abstract concepts and the following chart demonstrates the included terms as well as the repetition of them within the corpus:

Chart 5.3 Frequency of the abstract concepts

Relative Frequency of Abstract Concepts

A bar chart titled "Relative Frequency of Abstract Concepts" with a y-axis ranging from 0 to 1600 in increments of 200. The bars from left to right represent: way (~1380), life (~1030), home (~910), people (~860), love (~720), world (~590), mind (~425).

Diaspora has always been associated with abstract ideas including home, identity and in-betweenness and the frequent words such as *way, life, home, people, love, world* and *mind* have confirmed the significance of the abstract ideologies in diaspora. The analysis of the concordance lines and the collocations of the abstract ideas within the corpus showed that individuals in diaspora seek to express or convey their ideas, experiences and opinions about these abstract concepts. Also, it has been concluded that there is a consensus about the concept of home that it is a state of being, feeling and a quest for forming an identity and belonging, rather than being a physical entity and that is why it is mostly accompanied with the concept of mother because she strengthens the feeling of home.

The lowest state within the abstract concepts belongs to *mind,* which represents the rational thinking. The visualization of the corpus also unravels interesting information about

abstract concepts within the corpus. It was shown that the concepts of *love, home, way* and *life* are distributed closely to each other within the corpus. However, toward the end, it is the idea of *home* and *life* that remain closest to each other representing the necessity of finding a home to be able to live.

The evaluation of the corpus, in regard to objects and locations, reflected the fact that diasporic literature is very detail oriented when it comes to objects and objects are explained and addressed with certain elegance. The following chart also indicates the distribution of the terms that suggest the frequency of the objects and locations in the corpus :

Chart 5.4 Frequency of the locations and objects

Relative Frequency of Location and Object Nouns
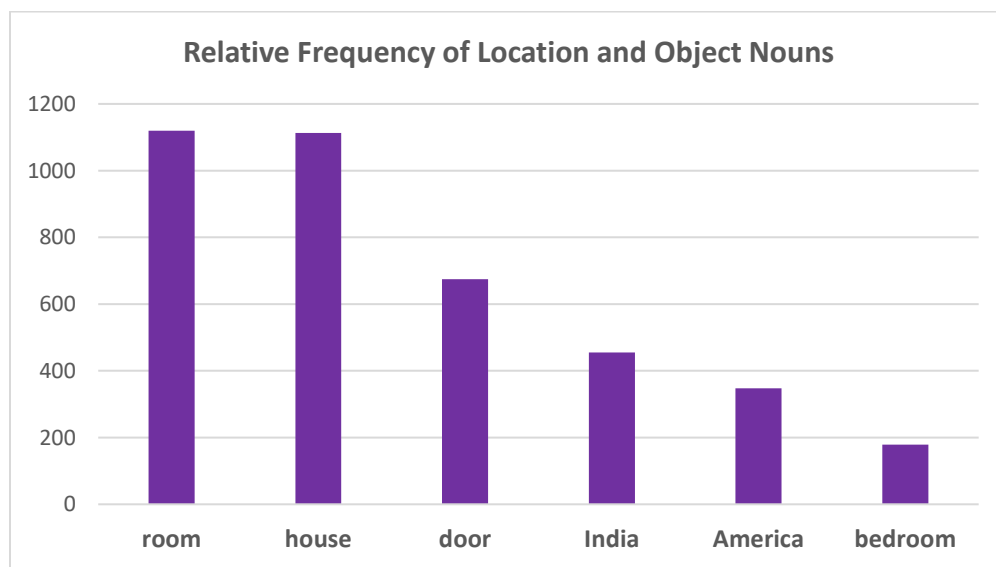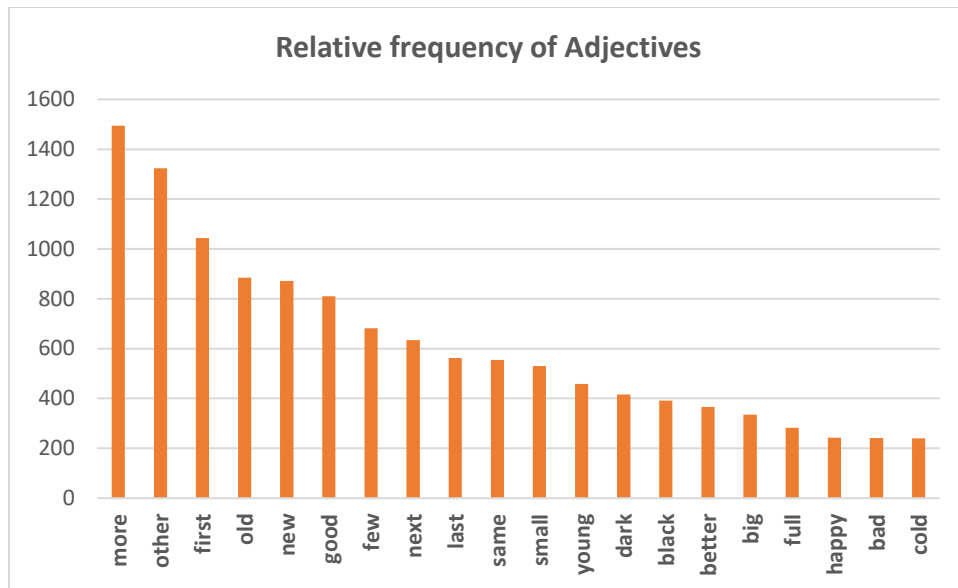
As the statistics of the most frequent terms in regard to objects indicate, words including *room, house, door* and *bedroom* are the most important objects that are mentioned in the corpus of the South Asian American literature and once again, they signify the importance of the concept of home as a safe resort. The examination of the references to the locations especially

about India and Indians showed that there is a tendency to extend the features that are mentioned about a type or group of Indians to all of them, for instance *all Indian parents, all Indian servants.* The only adjectives used prior to Indian even referring to their sales are *aggressive Indian sale* and *stupid Indian habit*.

Considering that diaspora provides the space for individuals to express their experiences, feelings, concerns and ideas both about the homeland and the host land , the evaluations of the most frequent verbs showed that verbs are mostly concerned with narrative verbs that give the tool to people to voice their experience and share them with the others. Being in-between gives the chance to the individuals to speak and that is why *said* is the most frequent verb of the corpus, followed by *talk, ask, tell* and *call*. Only four verbs are active verbs that include *go, put, come* and *make* and the rest are stative verbs such as *see, know, think, feel, look, seem, let, like and want*.

Within the adjective category, *more* with relative frequency of 1,496 is the most frequent one. *New* and *old* are two concepts that are common in diasporic literature in reference to home, culture, identity and setting, which are also among the most frequent adjectives of the corpus. Adjectives could be classified in almost three categories; descriptive, comparative and quantitative. Descriptive adjectives include *small, young, dark, black, big, full, happy, old, new, bad, cold* and *good* that explain the state of the being of a person or object. Comparative adjectives include *more*, *same* and *better* and quantitative adjectives contain *first, next, last* and *few*. Following is the chart of the most common adjectives within the corpus of this study:

Relative frequency of Adjectives

Bar chart showing relative frequency of adjectives (in descending order): more (~1490), other (~1330), first (~1040), old (~880), new (~870), good (~810), few (~680), next (~635), last (~565), same (~555), small (~525), young (~455), dark (~415), black (~390), better (~365), big (~335), full (~285), happy (~245), bad (~240), cold (~235).

5.2.1 The Comparison of the Findings with the Traditional Literary Analysis

One of the main questions that this research has attempted to answer was whether the study of diaspora through corpus tools and software adds to the existing literary readings of the diasporic literature or the findings dismiss the critical perspectives that are embedded in the literary theories.

It was argued that the contemporary perspectives on diaspora are mostly concerned with the domination of the Oriental discourse that is believed to be employed by the United States to categorize the people into the ethnical minority groups that they belong to. It was stated that once the people were categorized into the ethnical minority groups, the main features and characteristics that are associated with them also gradually turn into stereotypes about that minority group. The stereotypes that are associated with the South Asian American people were expressed afterwards and it was stated that these stereotypes were extended to the Diasporic literature as well. Here, the focus is on the diasporic literature written by the female authors and it was mentioned that authors in general had three approaches to these stereotypes. In the first

type, authors addressed and underlined the employed methods by the US to stabilize the Oriental doctrine and portray the South Asians as the inferiors. In the second type, authors turned the focus and the analysis of their novels into their own culture and challenged it and addressed the problems that exist within their own culture. Finally, the third approach is a multi-stream approach in which neither of cultures is in superior position and they considered diaspora as an organic being constituted of the different elements and social relations of both home and host lands.

All the evaluations that were carried out in regard to the evaluation of the structure of the corpus, wordlist, frequency list, concordance lines, collocations, keywords in context and visualization of the results signified the fact that the compiled corpus of the South Asian American Diasporic literature , written by female authors, actually follow the second approach of the mentioned perspectives. It was concluded that the focus of the corpus is not on the relation of the individuals with the host culture or country, but rather the emphasis is on the challenges and problems that exist within the culture of the homeland that people carry with them. Therefore, it could be argued that the South Asian American Diasporic literature has a critical stance against the traditional ideas and doctrines that are associated with the motherland.

On the other hand, the results of this research confirmed that women are in charge of carrying the culture and the values of the homeland in Diaspora and as the analytical perspectives addressed, the results of this study also confirmed that the suffering that women have to endure in diaspora is double because of both living in-between as well as bearing the patriarchal structure of their culture.

Also, as the results of this study have shown, the focus of the South Asian American Diasporic literature is on the challenges that the negative doctrines of the homeland cause, and

therefore, contrary to the literary viewpoints, there is no significant reference to the characters' intention to mimic the actions, behaviors and thoughts of the host land people. Moreover, unlike the existing theories on diasporic literature, the results of this study showed that there is no significant indication of the communication of those in diaspora with the outside world and with the people of the host land and it is the individual, interpersonal and the inner community relations that are focused and highlighted.

5.3 Suggestion for Further Studies

Given that this study was an attempt to uncover the main features of the diasporic literature through the employment of the quantitative and analytical tools, therefore the results of this study set the ground for future research concerning special concepts within diaspora including the representation of the language, identity, mimicry and culture.

Also considering the fact that the focus of this study was on the South Asian American female authors and the obtained results highlight the characteristics of the diasporic literature of this group, therefore a further study of the South Asian American diasporic literature written by male authors will provide a comprehensive view of the South Asian American Diaspora as a whole. At the same time, it can also indicate the similarities and the differences of the perspectives, approaches, ideas and writing style of the two groups.

In addition, considering that memory plays a significant role in keeping the connection with the old culture and identity within diaspora, it is believed that a detailed analysis of the portrayal of the memory within South Asian American Diasporic literature will offer significant findings about the role of memory in the protection of the image of the home. Corpus and

discourse analysis tools could be employed to examine how memory resonates home in actions, speeches and thoughts across the corpus.

Finally, the corpus analysis tools could be employed to identify the main structure as well as the features of the double consciousness in respect to women in diaspora and uncover the pattern of the double consciousness that is not identified by the literary analysis yet.

References

Abrams, M.H, and Geoffrey Galt Harpham. *A Glossary of Literary Terms*. McPeek, 1999.

---. *A Glossary of Literary Terms*. Wadsworth Cengage Learning, 2015.

Adur, Shweta M., and Bandana Purkayastha. "(Re)Telling Traditions: the Language of Social Identity among Queer South Asians in the United States." *South Asian Diaspora*, vol. 9, no. 1, 2016, pp. 1–16., doi:10.1080/19438192.2016.1199456

Ages, Arnold. *The Diaspora Dimension*. Springer, 1973.

---. "The Diaspora: Origin and Meaning." *The Diaspora Dimension*. Springer, Dordrecht,1973

Ang, Ien. "Together-in-Difference: beyond Diaspora, into Hybridity." *Asian Studies Review*, vol. 27, no. 2, 2003, pp. 141–154., doi:10.1080/10357820308713372.

Ashcroft, Bill, and D. P. S. Ahluwalia. *Edward Said: the Paradox of Identity*. Routledge, 2008.

---. *Edward Said*. Routledge, 2009.

Ashcroft, Bill, et al. *The Empire Writes Back: Theory and Practice in Post-Colonial Literature*. Routledge, 1994.

Badruddoja, Roksana. "Third World Woman," Family, and Marriage: South Asian Diasporic Fiction as a Site for Consolidation of the American Nation-state, South AsianReview, vol 35, no 2, 81-103, Manhattan College , DOI: 10.1080/02759527.2014.11932972

Bald, Vivek. *The Sun Never Sets South Asian Migrants in an Age of U.S. Power*. NYU Press, 2013.

Barlow, M. 2000. MonoConc Pro (Version 2.2) [Computer Software]. Available from http://www.athel.com/mono.html.

Balossi, Giuseppina. *A Corpus Linguistic Approach to Literary Language and Characterization: Virginia Woolf's "The Waves"*. Benjamins, 2014.

Baker, Paul. *Using Corpora to Analyze Gender*. Bloomsbury Academic, 2014.

---. *Using Corpora in Discourse Analysis*. Bloomsbury, 2013.

---. *Using Corpora to Analyze Gender*. Bloomsbury Academic, 2014.

---. *Sociolinguistics and Corpus Linguistics*. Edinburgh University Press, 2010.

Baumann, Martin. "Diaspora: Genealogies Of Semantics And Transcultural Comparison."
    Numen, vol. 47, no. 3, 2000, pp. 313–337., doi:10.1163/156852700511577.

Bhabha, Homi. *The Location of Culture*. Routledge, 1994.

---. "The Third Space: Interview with Homi Bhabha." *Identity: Community, Culture,*
    Difference, by Jonathan Rutherford, Lawrence & Wishart, 1998.

---. *Nation and Narration*. Routledge, Taylor & Francis Group, 2013.

Bhalla, Tamara Ayesha. "Between History and Identity: Reading the Authentic in South Asian
    Diasporic Literature and Community." *University of Michigan*, 2008.

Bhatia, Sunil. *American Karma: Race, Culture, and Identity in the Indian Diaspora*. New York
    University Press, 2007.

Bianchi , Francesca. "Corpora and Corpus Linguistics." *CULTURE, CORPORA AND*
    SEMANTICS Methodological Issues in Using Elicited and Corpus Data for Cultural
    Comparison, Universita Dei Siena , 2012, pp. 31–53.

Biber, Douglas, et al. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge
    Univ. Press, 2007.

Brah, Avtar. *Cartographies of Diaspora: Contesting Identities*. London, Routledge, 1996.

Brown, Garrett W.., et al. *The Concise Oxford Dictionary of Politics and International Relations*.
    OUP Oxford, 2018.

Brubaker, Rogers. "The 'Diaspora' Diaspora." *Ethnic and Racial Studies*, vol. 28, no. 1, 1 Jan.
    2005, pp. 1–19., doi:10.1080/759053713.

Burrows, J. F. *Computation into Criticism: a Study of Jane Austen's Novels and an Experiment*
    in Method. Clarendon Press, 1987.

Burman, Jenny. "Diasporic Pasts and Futures: Transnational Cultural Studies in Canada."
    TOPIA: Canadian Journal of Cultural Studies, vol. 17, 2007, pp. 5–10.,
    doi:10.3138/topia.17.5.

Burdick, Anne. *Digital_humanities*. Mit Press, 2016.

Chen, Kuan-Hsing. *Asia as Method: toward Deimperialization*. Duke University Press, 2010.

Chaudhry, L. (1998). "We are graceful swans who can also be crows" Hybrid identities of
    Pakistani Muslim women. In S.D. Dasgupta (Eds.), *A Patchwork Shawl: Chronicles of
    South Asian Women in America,* pp. 46-61, New Brunswick, Rutgers University Press,
    1989.

Chomsky, Noam. *Syntactic Structures*. Martino, 2015.

Cho, Lily. "The Turn to Diaspora." *Canadian Journal of Cultural Studies*, no. 17, 2007.

Chuang, Angie, and Robin Chin Roemer. "The Immigrant Muslim American at the Boundary of
    Insider and Outsider." *Journalism & Mass Communication Quarterly*, vol. 90, no. 1,
    2013,   pp. 89–107., doi:10.1177/1077699012468740.

Clifford, James. "Diasporas." *Cultural Anthropology*, Further Inflections: Toward Ethnographies
    of the Future, vol. 9, no. 3, 1994, pp. 302–338., JSTOR, www.jstor.org/stable/656365.

---. *Routes: Travel and Translation in the Late Twentieth Century*. London, Harvard
    University Press, 1997.

Cohen, Robin. *Global Diasporas: an Introduction*. Routledge, 2010.

---. *Global Diasporas: an Introduction*. 2nd ed., Routledge, 2008.

---. *Global Diasporas: an Introduction*. 2nd ed., University of Washington Press, 2008.

Cornis-Pope, Marcel. *New Literary Hybrids in the Age of Multimedia Expression: Crossing
    Borders, Crossing Genres*. John Benjamins Publishing Company, 2014.

Culpeper, Jonathan. "Keyness: Words, Parts-of-Speech and Semantic Categories in the
    Character-Talk of Shakespeare'SRomeo and Juliet." *International Journal of Corpus
    Linguistics*, vol. 14, no. 1, 2009, pp. 29–59., doi:10.1075/ijcl.14.1.03cul.

Dasgupta, Shamita Das. *A Patchwork Shawl: Chronicles of South Asian Women in America*.
    Rutgers Univ. Press, 2002.

Dash, Niladri Sekhar. "Context and Contextual Word Meaning." *SKASE Journal of Theoretical
    Linguistics*, vol. 5, no. 2, 2008, pp. 21–31.

Francis, Winthrop Nelson., et al. *Frequency Analysis of English Usage: Lexicon and Grammar*. Houghton Mifflin, 1982.

Gallagher, Michael David. "Producing the Schooled Subject: Techniques of Power in a Primary School Classroom." Institute of Geography, University of Edinburgh, 2004.

Gandhi , Leela. *Postcolonial Theory A Critical Introduction*. Columbia University Press, 1998.

Gardiner, Eileen, and Ronald G. Musto. *The Digital Humanities: a Primer for Students and Scholars*. Cambridge University Press, 2015.

Ghosh, Ranjan. *(In)Fusion Approach: Theory, Contestation, Limits: (in)Fusionising a Few Indian English Novels*. Maryland, University Press of America, 2006.

Girishkumar , Divya. "Diaspora and Multiculturalism: British South Asian Women's Writing." *Cardif University*, 2014.

Gibbs, Raymond W. *The Poetics of Mind: Figurative Thought, Language, and Understanding*. Cambridge University Press, 2010.

Gold, Matthew K., and Lauren F. Klein. *Debates in the Digital Humanities: 2016*. University of Minnesota Press.

Goldberg, David Theo. "Deprovincializing Digital Humanities." *Between Humanities and the Digital*, by Patrik Svensson and David Theo Goldberg, The MIT Press, 2015, pp. 163–173.

Grewal, I. "Reading and Writing the South Asian Diaspora: Feminism and Nationalism in North America." In *Our Feet Walk the Sky: Women of the South Asian Diaspora*, Aunt Lute Books, 1993.

Hall, Stuart. "Cultural Identity and Diaspora." *Colonial Discourse and Post-Colonial Theory: A Reader*, edited by Laura Chrisman, Patrick Williams, Harvester Wheatsheaf, 1994, pp. 227–237.

---. "Cultural Identity and Diaspora." *Colonial Discourse and Post-Colonial Theory A Reader*, by Patrick Williams and Laura Chrisman, New York ,Colombia University Press, 1994, pp. 392–403.

Herold, Mariella Espinoza, and Rina Manuela. Contini. *Living in Two Homes. Integration and Education of Transnational Migrants in a Globalized World*. Emerald Group Publishing, 2017.

Hockey, Susan M. *Electronic Texts in the Humanities: Principles and Practice*. Oxford
University Press, 2014.

Hoogvelt, Ankie Maria Margaretha. *Globalization and the Postcolonial World: the New Political
Economy of Development*. Palgrave Macmillan, 2007.

Hussain, Yasmin. *Writing diaspora: South Asian women, culture and ethnicity*. S. Husain (Ed.).
Voices of resistance: Muslim women on war, faith, & sexuality. Burlington, VT: Ashgate,
2005.

---. *Writing Diaspora; South Asian Women, Culture and Ethnicity*, Taylor Francis Ltd, United
Kingdom, 2005.

Hughes, Lorna M. *Evaluating and Measuring the Value, Use and Impact of Digital Collections*.
Facet, 2012.

Hori, Masahiro. "Collocation Style of First-Person Narratives in Dickenson: *David Copperfield*,
Bleak House and *Great Expectations* ." *Stylistic Studies of Literature: in Honour of
Professor Hiroyuki Itō*, by Masahiro Hori et al., Peter Lang, 2009.

Ide, Nancy. "Preparation and Analysis of Linguistic Corpora." *A Companion to Digital
Humanities*, pp. 289–305., doi:10.1002/9780470999875.ch21.

Jain, A. "Differences in Difference: A Cognitive Mapping of the Landscape of Otherness." Paper
presented at *Edward Said and Orientalism Conference*, Boston, 2005.

Jensen, Kim Ebensgaard. "Linguistics in the Digital Humanities: (Computational) Corpus
Linguistics." *MedieKultur: Journal of Media and Communication Research*, vol. 30, no.
57, 2014, p. 20., doi:10.7146/mediekultur.v30i57.15968.

Kalra, Virinder S., et al. *Diaspora & Hybridity*. SAGE, 2010.

Kenny, Kevin. *Diaspora: a Very Short Introduction*. Oxford University Press, 2013.

---. *Diaspora: a Very Short Introduction*. Oxford University Press, 2013.

Khadem, Amir. "Annexing the Unread: a Close Reading of 'Distant Reading.'" *Neohelicon*, vol.
39, no. 2, 2012, pp. 409–421., doi:10.1007/s11059-012-0152-y.

Lane, Richard J. *The Big Humanities: Digital Humanities/Digital Laboratories*. Routledge,
2016.

Layfield, Allison. "Asian American Literature and Reading Formations: A Case Study of Nora Okja Keller's *Comfort Woman* and *Fox Girl*." Reception: Texts, Readers, Audiences, History, vol. 7, no. 1, pp. 64–82. JSTOR, JSTOR www.jstor.org/stable/10.5325/reception.7.1.0064, 2015.

Leena Pundir " Globalization and Literature: Female South Asian Diasporic Voices", New Academia: An International Journal of English Language, Literature and Literary Theory, Vol. IV Issue I, Jan. 2015

Loomba, Ania. *Colonialismo/Postcolonialismo*. 2nd ed., Routledge , 2005.

---. *Colonialism/Postcolonialism*. Routledge, Taylor & Francis Group, 2015

Macfie, A.L. *ORIENTALISM*. London, Longman, 2002.

Mani, L "Gender, Class and Cultural Conflict: Indu Krishnan's Knowing Her Place." *In Our Feet Walk the Sky: Women of the South Asian Diaspora*, ed The Women of South Asian Descent Collective, pp 32-36, San Francisco, CA: Aunt Lute Books, 1993.

Malhi, Rebecca L., et al. "`Being Canadian' and `Being Indian': Subject Positions and Discourses Used in South Asian-Canadian Women's Talk about Ethnic Identity." *Culture & Psychology*, vol. 15, no. 2, 2009, pp. 255–283.

Mcclintock, Anne. "The Angel of Progress: Pitfalls of the Term 'Post-Colonialism." *Social Text*, no 31/32, 1 Jan. 1992, www.jstor.org/stable/466219.

McLean, Iain, and Alistair McMillan. *The Concise Oxford Dictionary of Politics*. Oxford University Press, 2009.

Mitchell, Katharyne. "Different Diasporas and the Hype of Hybridity." *SAGE Journals*, vol. 15, no. 5, 1997, pp. 533–553.

Mohanram, Radhika*, Black Body, Women, Colonialism, and Space,* Minneapolis, University of Minnesota Press, 1999.

Moretti, Franco. "Conjectures on World Literature." *New Left Review*, vol. 1, no. 4, 2014, pp. 54–68.

Naber, Nadine. "Arab American Femininities: Beyond Arab Virgin/ American(Ized) Whore." *Feminist Studies*, vol. 32, no. 1, 2006, pp. 87–111. *JSTOR*, JSTOR, www.jstor.org/stable/20459071.

Nelson, Cary and Lawrence Grossberg ,*Marxism and the Interpretation of Culture,* pp. 271-313, Urbana, IL, University of Illinois Press, 1987.

Ngo, Bic,zand Stacey J. Lee. "Complicating the Image of Model Minority Success: A Review of Southeast Asian American Education." *Review of Educational Research*, vol. 77, no. 4, 2007, pp. 415–453., www.jstor.org/stable/4624907.

Oonk, Gijsbert. *Global Indian Diasporas: Exploring Trajectories of Migration and Theory*. Amsterdam University Press, 2011.

Quayson, Ato. *Postcolonialism: Theory, Practice or Process?* Polity Press, 2005.

Oxford English Dictionary, 1999, 10th edition, Judy Pearsall (ed.). New York, Oxford University Press.

Pang, Valerie Ooka, et al. "Asian American and Pacific Islander Students: Equity and the Achievement Gap." Educational Researcher, vol. 40, no. 8, 2011, pp. 378–389. JSTOR, JSTOR, www.jstor.org/stable/41302978.

Portes, Alejandro. "Immigration Theory for a New Century: Some Problems and Opportunities." *The International Migration Review*, vol. 31, no. 4, 1997, pp. 799–825. *JSTOR*, JSTOR, www.jstor.org/stable/2547415.

Puar, Jasbir K. "Writing my way "home": Travelling South Asian bodies and Diasporic Journeys." Socialist Review, 1994, vol 24, no 4, 75-108.

Ramsay, Stephen. *Reading Machines: toward an Algorithmic Criticism*. University of Illinois Press, 2012.

Ramsay, Stephen. *Reading Machines: Toward an Algorithmic Criticism*. University of Illinois Press, 2011.

Rauscher, Janneke. "Grasping Cities through Literary Representations. A Mix of Qualitative and Quantitative Approaches to Analyze Crime Novels." *Historical Social Research / Historische Sozialforschung*, vol. 39, no. 2 (148), 2014, pp. 68–102., www.jstor.org/stable/24142683.

Rangahau, Te Oru. "Hybridity in the Third Space : Rethinking Bi-cultural Politics in Aotearoa / New Zealand." Maori Research and Development conference, 7-9 July 1998, Massey University,2000.

Replogle, Elaine, and Sunaina Marr Maira. "Desis in the House: Indian American Youth Culture in New York City." *Contemporary Sociology*, vol. 32, no. 2, 2003, p. 194., doi:10.2307/3089591.

Rettberg, Scott. *Electronic Literature*. Polity Press, 2018.

---. "Electronic Literature as Digital Humanities." *A New Companion to Digital   Humanities*, by Susan Schreibman et al., John Wiley & Sons, 2016.

Rommel , Thomas. "Literary Studies ." *A Companion to Digital Humanities*, by Susan Schreibman, Blackwell Publ., 2011.

Said, Edward W. *Orientalism*. New York,Vintage Books, 1979.

---. *Orientalism*. W. Ross MacDonald School, Resource Services Library, 2006.

---. *Reflections on Exile: and Other Literary and Cultural Essays*. Granta, 2012.

---. *Reflections on Exile and Other Essays*. Granta Books, 2012.

---. *Culture and Imperialism*. Vintage Digital, 2014.

Salam, R. A, "Am I South Asian? Personal, organizational, and political understandings of second generation South Asians' Ethnic Identity*"*, Paper presented at the American Sociological Association ,2005

Saran, Rupam. "Model minority imaging in New York: The situation with second generation Asian Indian learners in middle and secondary schools." *Anthropologist Special Issue, 2*, 67-79, 2007.

Schreibman, Susan. *A Companion to Digital Humanities*. Blackwell Publ., 2011.

Schreibman, Susan, et al. *A Companion to Digital Humanities*. John Wiley & Sons, 2008.

---. *A New Companion to Digital Humanities*. John Wiley & Sons, 2016.

---. *The Digital Humanities and Humanities Computing: An Introduction*. Oxford: Blackwell, 2004.

Shah, Amita Roy. " Exploring Hybrid Identities: South Asian American Women Pursue a Career in Teaching", ProQuest LLC, Ed.D. Dissertation, Columbia University, 2013

Spivak, G. "Can the subaltern speak?", in C. Nelson & L. Grossberg (Eds.), *Marxism and the Interpretation of Culture*, Urbana, University of Illinois Press, pp. 271–313, 1988.

Sriram, Shyam Krishnan, and Stonegarden Grindlife. "The Politics of Deracialisation: South Asian American Candidates, Nicknames, and Campaign Strategies." *South Asian Diaspora*, vol. 9, no. 1, 2016, pp. 17–31., doi:10.1080/19438192.2016.1199503.

Stubbs, Michael. "Conrad in the Computer: Examples of Quantitative Stylistic Methods." *Language and Literature*, vol. 14, no. 1, 2005, pp. 5–24.

Teranishi, R., & Tchen, J. K. W. "Asian Americans and Pacific Islanders: Facts, not fiction: setting the record straight", CARE: National Commission on Asian American and Pacific Islander Research in Education, College Board, 2008.

Thomas, William G. "The Promise of the Digital Humanities and the Contested Nature of Digital Scholarship." *A New Companion to Digital Humanities*, 2015, pp. 524–537., doi:10.1002/9781118680605.ch36.

Tognini-Bonelli, Elena. *Corpus Linguistics at Work*. John Benjamins Publishing Company, 2001.

Toloyan, Khachig,.*Diaspora: a Journal of Transnational Studies, 1991*

---. "Diaspora Studies Past, Present and Promise", Paper presented at the Oxford Diasporas Programme, June 2011.

---. "The Nation-State and Its Others: In Lieu of a Preface." *Diaspora: A Journal of Transnational Studies*, vol. 1 no. 1, 1991, pp. 3-7. Project MUSE, doi:10.1353/dsp.1991.0008.

Um, Hae-Kyung. *Diasporas and Interculturalism in Asian Performing Arts: Translating Traditions*. Routledge, 2006.

Williams, Misti L. "Diasporic Awakenings: Mobilizing the Diaspora for Homeland Causes." University of Washington , 2010.

Young, Robert. *White Methodologies*. Routledge , 2004.

Zebroski , James Thomas. "Theory in the Diaspora ." *Jac*, vol. 24, no. 4, 2005.

Appendices

Appendix i

List of TagAnt Tag Sets

POS Tag	Description
CC	coordinating conjunction
CD	cardinal number
DT	determiner
EX	existential there
FW	foreign word
IN	preposition/subord. conj.
IN/that	complementizer
JJ	adjective
JJR	adjective, comparative
JJS	adjective, superlative
LS	list marker
MD	modal
NN	noun, singular or mass
NNS	noun plural
NP	proper noun, singular
NPS	proper noun, plural
PDT	predeterminer
POS	possessive ending
PP	personal pronoun
PP$	possessive pronoun
RB	adverb
RBR	adverb, comparative
RBS	adverb, superlative
RP	particle
SENT	end punctuation
SYM	symbol
TO	*to*
UH	interjection
VB	verb *be*, base form
VBD	verb *be*, past
VBG	verb *be*, gerund/participle

VBN	verb *be*, past participle
VBZ	verb *be*, pres, 3rd p. sing
VBP	verb be, pres non-3rd p.
VD	verb do, base form
VDD	verb do, past
VDG	verb do gerund/participle
VDN	verb do, past participle
VDZ	verb do, pres, 3rd per.sing
VDP	verb do, pres, non-3rd per.
VH	verb have, base form
VHD	verb have, past
VHG	verb have, gerund/participle
VHN	verb have, past participle
VHZ	verb have, pres 3rd per.sing
VHP	verb have, pres non-3rd per.
VV	verb, base form
VVD	verb, past tense
VVG	verb, gerund/participle
VVN	verb, past participle
VVP	verb, present, non-3rd p.
VVZ	verb, present 3d p. sing.
WDT	wh-determiner
WP	wh-pronoun
WP$	possessive wh-pronoun
WRB	wh-abverb
:	general joiner
$	currency symbol

Appendix ii

The frequency list of the corpus

N	Word	Raw Frequency	Relative frequency	%
1	Said	9009	4267	0.25
2	TIME	4050	1917	0.19
3	MOTHER	3300	1519	0.16
4	EYES	2713	1285	0.13
5	DAY	2394	1127	0.11
6	HOUSE	2358	1114	0.11
7	FATHER	2349	1071	0.11
8	LIFE	2200	1035	0.10
9	LOOKED	2178	1032	0.10
10	HEAD	2160	1019	0.10
11	HAND	2114	993	0.10
12	TOLD	2105	997	0.10
13	HOME	1932	911	0.09
14	TELL	1924	911	0.09
15	LOOK	1896	890	0.09
16	PEOPLE	1835	862	0.09
17	FAMILY	1780	837	0.08
18	LITTLE	1720	815	0.08
19	LEFT	1578	747	0.07
20	LOVE	1524	721	0.07
21	RHE	1459	691	0.07
22	WOMAN	1458	681	0.07
23	NIGHT	1447	680	0.07
24	DOOR	1427	675	0.07
25	FEROZA	1350	639	0.06
26	VOICE	1392	659	0.07
27	HANDS	1373	650	0.06
28	HAIR	1273	601	0.06
29	PADMA	1269	601	0.06
30	WORLD	1254	586	0.06

31	PARENTS	1244	589	0.06
32	HEARD	1213	574	0.06
33	NAME	1193	565	0.06
34	BED	1189	556	0.06
35	LOOKING	1183	560	0.06
36	CHILD	1182	548	0.06
37	DAYS	1169	553	0.06
38	BOY	1157	530	0.05
39	SAT	1126	532	0.05
40	WATER	1105	518	0.05
41	YES	1087	513	0.05
42	FOUND	1044	494	0.05
43	WOMEN	1043	489	0.05
44	CHILDREN	1028	479	0.05
45	BODY	1019	482	0.05
46	INDIA	977	456	0.05
47	HUSBAND	968	458	0.05
48	CALLED	963	456	0.05
49	FEEL	963	452	0.05
50	INSIDE	963	456	0.05
51	WHITE	961	455	0.05
52	WIFE	945	439	0.04
53	GIRL	935	443	0.04
54	MORNING	930	439	0.04
55	MARRIED	929	441	0.04
56	WORDS	926	586	0.04
57	GONE	913	432	0.04
58	SON	911	431	0.04
59	MIND	905	428	0.04
60	SEEN	900	426	0.04
61	CAR	894	423	0.04
62	STOOD	886	419	0.04
63	TALK	880	411	0.04
64	DARK	878	411	0.04
65	READ	872	409	0.04
66	ING	866	407	0.04
67	INDIAN	864	408	0.04
68	COURSE	863	408	0.04
69	FRANK	851	403	0.04

70	CALL	838	397	0.04
71	SCHOOL	832	390	0.04
72	TABLE	828	391	0.04
73	BLACK	826	389	0.04
74	AMERICA	825	387	0.04
75	MA	824	390	0.04
76	LIGHT	822	389	0.04
77	DIDN	821	389	0.04
78	DAUGHTER	817	382	0.04
79	WALKED	802	380	0.04
80	OH	801	378	0.04
81	MOMENT	800	378	0.04
82	KITCHEN	798	378	0.04
83	HELP	796	376	0.04
84	ELSE	785	372	0.04
85	LEAVE	780	369	0.04
86	FRIENDS	779	369	0.04
87	HEART	773	366	0.04
88	RAZA	752	356	0.04
89	RED	748	352	0.04
90	HELD	747	354	0.04
91	TRIED	746	353	0.04
92	SET	744	350	0.04
93	SAYING	743	352	0.03
94	AMERICAN	742	348	0.03
95	HEAR	740	350	0.03
96	FRONT	739	350	0.03
97	PAST	739	350	0.03
98	MONEY	731	353	0.03
99	STOP	713	325	0.03
100	PREMA	709	336	0.03