

2. Literature Review

In order to carry out traditional CDA research from a novel perspective, while additionally proposing and testing a new method through which to carry out future CDA-based investigations, it is first important to discuss the field itself. As a result, the following section highlights the history of CDA, its development, and some of its applications within the literature. Further, this section discusses the problems which have been found in the traditional CDA approach to linguistics research as well as the linguistic research methods which were used in the current study as potential solutions to pre-existing methodological weaknesses.

With the intention of laying the groundwork for the investigation which was undertaken here, this section reviews past and present research in all of the fields included in the present methodology. This is important not only in clarifying the logical steps which have influenced the current research but in presenting the methodologies being combined here in terms of how they can serve to solve some common problems found in traditional CDA research. As such, all of the research methods discussed below are looked at with a basis in their connections (both explicit and implicit) to CDA research as well as in their historical presence and achievements in various fields of linguistics research.

2.1. Critical Discourse Analysis

CDA—as was discussed in the introduction (see page 1)—is an approach to linguistics research which differs from classical DA, “...in not just describing discursive practices, but also showing how discourse is shaped by relations of power and ideologies, and the constructive effects discourse has upon social identities, social relations and systems of

knowledge and belief, neither of which is normally apparent to discourse participants,” (Fairclough, 1992, p. 12). Although CDA has been applied to a great number of issues within linguistics research, all while using many different methodological approaches, it is safe to say that Fairclough’s description encapsulates the basest characteristics which set CDA apart from other areas of linguistics (most notably DA) and also what has caused there to be such debate in the field over what CDA is, where it is going and whether it plays a legitimate role in advancing linguistics research as a whole.

One of the most difficult aspects of discussing CDA is actually pinning the area down in concrete terms. The area, or methodology, or theory (depending on the author consulted⁴) can have a number of uses, from political speeches, to racism, to advertising; and not surprisingly, the history of CDA is similarly murky and dependent on the perspective of each author. What is generally agreed upon in the field is that CDA came about sometime between the mid-1970s and early 1980s as a response to both DA, which came to prominence in the 1960s (van Dijk, 1997), and critical linguistics (Wodak & Meyer, 2009; Fairclough, 1995).

In its simplest form, CDA research can be conceptualized as a pyramid representing the multi-directional interaction between discourse, cognition, and society (van Dijk, 1997). In this model, discourse is seen to be a socially and ideologically based cycle where ideologies are seen as, “...representations of aspects of the world which can be shown to contribute to establishing, maintaining and changing social relations of power, domination and exploitation,” (Fairclough, 2003, p. 9). Due to this perceived connection between

⁴ In this work, CDA is referred to as an *area* or *field* (in reference to the body of work of authors working in CDA) as well as either an *approach* or a *methodology* (referring to CDA’s use in carrying out a given study or set of studies).

discourse, 'power,' 'domination' and 'exploitation,' researchers' ideological leanings and understanding of modern and historical events become important in interpreting discourses as well as in selecting data for studies. For example, in looking at political discourses one must account for political knowledge, underlying political strategies and ideological stances of the originator of the discourse being examined, as well as numerous sociological and cultural factors in order to render a suitable analysis which addresses discourse, cognition, and society (van Dijk, 1997). In order to effectively address these parts of the discourse, it is necessary to utilize a methodology which is focused on looking at language use through the lens of ideological and social factors, and, indeed, this is what has been done in much of the published CDA literature to date. Despite differences of opinion in the area, the presence of 'discourse, cognition and society' can be seen in most studies. In Wodak and Weiss' (2003) four-level approach, for instance, CDA research is seen as a way to 'triangulate' discourses in context using four steps. First, the language or text is looked at descriptively. This is followed by an examination of the relationships between the 'utterances, texts, genres and discourses' found therein. Finally extralinguistic (socially based) variables are examined and these (as well as the discourse itself) are looked at in terms of 'sociopolitical and historical contexts.' The presence of the 'pyramid' is also found in other, less well-known approaches and forms the basis for nearly all CDA research.

It is because of the near-universality of the two underlying characteristics present in the models mentioned above (ideology and social factors) that there is a prevalence of interdisciplinary methodologies within most studies in the area. Consequently, these characteristics also contribute to many of the criticisms leveled against CDA, primarily that studies fall prey to 'circularity' and that they are often not based on 'firm linguistic

evidence' (Stubbs, 1997). If discourse is seen to affect not only individuals but also society and—as a consequence—future discourses, then understanding discourse and effectively utilizing CDA methods necessarily requires the use of varied discourse arenas. This is true methodologically speaking in that authors may borrow from other fields while conducting research; however, it also signifies that CDA is—for good or ill—one of the most flexible areas of linguistics research in terms of applicability. Because 'discourse' is present in so many areas of human existence and is, as a result, difficult to aptly define (see van Dijk, 1997), CDA can be carried out in many different linguistic areas. Since its inception, CDA has focused on a number of issues, while, "gender issues, issues of racism, media discourses, political discourses, organizational discourses or dimensions of identity research have become very prominent. The methodologies differ greatly in all these studies on account of the aims of the research and also the methodologies applied: small qualitative case studies can be found as well as large data corpora, drawn from fieldwork and ethnographic research," (Weiss & Wodak, 2003, p. 12). While there is a great deal of diversity in the applications that have been realized by CDA researchers, because of its reliance on ideology and social matters, much of the area's attention has been devoted to examining political and media discourses. The current study picks up where many CDA examinations of the media have left off by addressing methodological weaknesses and focusing on the media's presence in modern society.

2.2. CDA Studies of Media

In order to carry out the current research it is first important to examine CDA research as it is tied to studies of the mass media. This is the case in that the current project is an examination of media discourse using CDA as a starting point. Further, although there

have been media-centered studies using CDA in the past (see below) much of the published research on media language use has existed as a fundamentally separate entity from CDA. With this in mind, the first step in approaching a CDA-centric media investigation was to highlight the shared characteristics and complimentary methods between CDA research and general media studies.

Although the media is an enormous and prominent source of socially significant discourse and language use, the linguistic study of media is relatively undeveloped. Much of what is today thought of as Media Studies was initiated by the Glasgow Media Group under the direction of Greg Philo in the mid-1970s (Glasgow Media Group, 2010; Fürsich, 2009). Following initial forays into the study of the media's presence in society, the discipline began to grow, finally expanding to form a large, extremely interdisciplinary branch of DA. Eventually, with the introduction of research components related to Critical Linguistics in the late 1980s, Media Studies began to have a presence in CDA-based research as well. In this respect, CDA came to be used as a means to explore the symbiotic relationship between media and discourse construction. This understanding of discourse's place in the media relies on the idea that a society or culture's day-to-day reality influences what is presented in the media which, in turn, influences the same society or culture's day-to-day reality (Gutiérrez Vidrio, 2010). This is encapsulated quite well in Carvalho's (2008) concept of Mediated Discourse.

The combination of Media Studies and CDA came about as a natural progression due to CDA's simultaneous embrace of aspects of sociology and political science, as well as syntax and semantics. Due to CDA's focus on the discursive presence of divisive social events in modern society and the news media's reporting on such events, the combination

of CDA and Media Studies provides a doubly-representative view of modern critical linguistic issues. While CDA highlights the presence of discourse in media language and the potential consequences, Media Studies serves as a window into the prevalence of such language use in practical areas. Indeed, for these same reasons, numerous worthwhile studies have come out of critical analyses of media. Many of the studies carried out involving CDA as applied to Media Studies have focused on ‘framing’ (Túñez & Guevara, 2009, for example) and similar instances of what authors have called ‘imaging’ and ‘portrayal’ (see, e.g., Berger, 2009). These studies are important examples of the successful pairing of Media Studies and CDA because they show not only the underlying discourses present in language but common occurrences of such discourses.

When viewing the use of combined CDA and Media Studies methodologies, ‘framing’ studies would seem to be some of the most common approaches. These studies are based on the concept of discursive frames, which according to Lewis and Reese (2009) are,

a central facet of political communication, frames define the terms of debate; shape public opinion through the persuasive use of symbols; and, when most effective, lead to public policy change. They serve as the primary vehicle through which public officials, the news media, and other elites exercise political influence over one another and the public at large. (p. 85)

According to this conceptualization, frames carry out many of the same functions that CDA is charged with studying in relation to discourse; this is primarily true with regard to the idea that frames can affect ‘change’ in the public and private sectors. Because of this, framing studies represent one of the most common approaches within the discipline.

Because frame research in media language is so conceptually similar to the fundamental principles of CDA, the topics in which framing is examined are usually tied to socially important, controversial, or political themes. Lewis and Reese (2009), for example, studied framing in the media following the events of September 11th in the United States. The research team viewed the phrase ‘War on Terror’ as a frame which, to them, constituted a, “...socially shared organizing principle...” (p. 86). Their study examined this term’s use and rise to prominence within journalistic texts. In order to gain a perspective on its use as well as journalists’ relation to the framing of ‘War on Terror,’ Lewis and Reese carried out interviews with journalists in order to delineate journalism’s role in a similar way to that used in Carvalho’s (2008) Mediated Discourse. That is, a cycle in which journalists put forth a term (through publication) with a certain perspective—influenced by their day-to-day realities. The cycle which Carvalho proposes—and which influenced the current proposal of Methodological Fusion—involves a given term being taken up by the press (usually after first being emitted by another entity) with a certain context/perspective in mind; and finally the term is disseminated to the public who gleans a certain meaning from it. This meaning then goes on to influence further media coverage involving the term. Carvalho posits that this is true in that the connection between discourse and society means that, “each discursive event is dialectically tied to society insofar as it both constitutes and is constituted by social phenomena,” (p. 162). Versions of this sort of cycle are present in many areas of the published literature, and indeed form the very foundation of traditional CDA research. Oster (2010) (citing Kövecses, 1990) stresses that even something as complex as emotion can be described through lexical analysis in that it is present in language use through the use of metaphor and collocations (both of which are obviously tied to ‘social phenomena’ as well).

In one of the most transparent examples of Carvalho's (2008) process, the phrase 'War on Terror' was first put forth by politicians before being spread and given currency through its presence in news media where it has since become a very commonplace concept among the public (Lewis & Reese, 2009). In Carvalho's model, after a term gains currency through media permeation, the public's perspective progresses to the point that it affects policy at which point (in one way or another) the cycle begins anew. This can be seen in the same example of 'War on Terror' where the term went from politicians to journalists to the public before becoming a divisive ideological stance in American politics which has since entered the common lexicon and influenced numerous public and private policies.

Many other authors have used similar strategies to identify and understand framing in media discourse. Some of these studies have taken a more straightforward approach to the field, whereas others have viewed framing from a more abstract perspective than in studies like that carried out by Lewis and Reese (2009). Many of the studies which fall into the more theoretically concrete end of the spectrum have addressed 'imaging' and 'portrayal' in the media. Berger (2009), for instance, examined the palpable effects of media framing on a population in relation to disaster coverage in the media. By analyzing what Berger identified as the juxtaposition between the 'mediated representation' and the 'lived reality' of New Orleans following hurricane Katrina, the author examined the, "...merger of journalistic crisis and state breakdown [which] involved, among other things, mutually reinforcing notions of a crime-and-punishment spectacle that demonized much of New Orleans' population that weathered the storm," (p. 491). Thus, the Berger study gives credence to Carvalho's (2008) concept of Mediated Discourse by showing two sides of it. Berger's conceptualization of a 'representation' and a 'reality' within news coverage

corresponds with Carvalho's cycle in which the media portrays reality through reporting which is presented as unbiased and representative; this, in turn, influences public perception and policy.

In a similar vein, Davis and French (2008) looked at the connection between news and day-to-day life among New Orleans citizens following Hurricane Katrina. Focusing particularly on media characterizations of New Orleans citizens, Davis and French concluded that, "Post-Katrina news depictions were found to rely on specific rhetorical devices and semantic strategies, as well as seemingly value-free terminology (e.g., victim and survivor) to publicly discuss certain citizens involved in the event. However, such discursive construction had the rhetorical result of shifting blame onto those citizens," (p. 244). The difference between this study and the Berger (2009) investigation is that instead of viewing coverage as a means of portraying a different lived reality (as Berger did), Davis and French viewed framing in news coverage following Hurricane Katrina as a reflection of pre-existing frames regarding New Orleans. This is similar to the concept of Mediated Discourse (Carvalho, 2008) in that discourse's influence on society and life is viewed as a sort of cycle in which pre-existing discourses are important in shaping media discourses, which in turn influence lived reality and the perception of it. In the case of Davis and French's study, the frames examined essentially consisted of cultural stereotypes based around what the authors termed New Orleans' reputation as a city that is 'free-wheeling' and 'easy going.'

While the above studies focused on concrete examples of rather isolated uses of discourse in the media, other prominent investigations have focused on the more nebulous, philosophical side of media discourse studies by looking at individual fragments of news media language use as they relate to framing. These studies have included examinations of

media characteristics such as framing in leads (the opening sentences in newspaper articles) in internet news articles (Zillmann, Chen, Knobloch, and Callison, 2004) and the historical development of individual lexical items (LIs) within the media (see Peñaranda Contreras, 2010) according to what Berger (2009) calls ‘lived reality.’ Despite the fact that investigation of historical discourse transformation is quite common (see Peñaranda Contreras, 2010; Wodak, 2007; Törrönen, 2004) when analyzing the formulation and propagation of frames a less ample gamut of research approaches is present in the literature.

Nonetheless, researchers have looked at both the construction of frames and the development of news media reports. This has been done on both macro and micro levels. Túñez Guevara (2009), for example, looked at the presentation of one story through the print news media of four different locations (Mexico, Portugal, Spain, and Galicia). This approach allowed for the examination of contrasts between Mediated Discourses according to the media environment in which they were present. What was most noteworthy in this study was the way in which it was approached. By comparing media discourses according to country of publication—while controlling for topic—Túñez Guevara was able to single out differences and see them as something more significant than just the perception of discourse characteristics. That is, while a difference within a single country could be indicative of differing opinions or perspectives, international differences (and, by the same token, similarities) could be seen as indicators of differences in presentation and interpretation. This sort of control over the analysis of media discourse is what informed the methodological approach designed for the current study. If variables can be controlled well enough, and if the language present in the texts examined is ‘neutral’ enough, then—in theory—discourse characteristics will ‘rise to the top’ based on their mere presence in the texts analyzed.

This same sort of methodologically controlled analysis has also been carried out on a micro level, Van Hout and MacGilchrist (2010) ‘followed’ a news story through the publication process in order to observe and analyze, “...how discursive transformations shed light on journalists’ writing practices... [they also] argue that the notion of framing should be decoupled from the intent to mislead; [because] every news story, as indeed every type of linguistic expression, must *per se* be ‘framed,’” (p. 170). What this means is that Carvalho’s (2008) concept of Mediated Discourse has such an effect on discourse formation and presentation that it becomes impossible to encounter media discourse that is not influenced by social and political factors in some way. This outlook on media discourse has been supported through other examinations of media language use based on Wodak’s DHA (see Peñaranda Contreras, 2010; Törrönen, 2004).

These sorts of studies are not as common in the literature as are frame-based studies. Nonetheless, in the same way that studies like Van Hout and MacGilchrist (2010) have examined the micro-development of frames, authors such as Peñaranda Contreras (2010) and Törrönen (2004) have analyzed the macro-presence of discourse constructions as well as the changes that these constructions incurred over time; in the case of the Peñaranda Contreras study, the research centered on the changing use of lexical forms in news reporting during the 1960s and 1970s. The study was carried out using Colombian newspapers from two decades which were analyzed for the way in which they referred to those involved in the drug trade. The study found that over time, due to changes in day-to-day life in the country as well as a general shift in Colombia’s lived reality, newspapers shifted their use from the term ‘marihuaneros’ (loosely translated by the author as *drug fiends*—employing American colloquial speech, one could even go so far as to suggest *pot heads*) to ‘mafiosos’ (translated by the author as *drug lords*). Similarly, Törrönen (2004)

analyzed articles about drugs and drug use viewing the rhetoric employed therein through the political context in which it occurred. Törrönen found similar changes according to the overall public and political opinion of the issue at a given time. In both of these studies, the most important binding characteristic is their focus on discourse and Mediated Discourse (Carvalho, 2008). Although both studies utilized methodologies similar to those present in some of the preceding discussion, the brunt of their arguments seems to be that discourse can be seen in a society's day-to-day 'reality' and that a society's day-to-day 'reality' can similarly be seen in discourse use.

The one negative characteristic which these studies share (at least in terms of the CDA weaknesses addressed in the present work), is a basis in the liberal selection of texts and parts of text for analysis. This methodological characteristic is certainly not relegated to the Peñaranda Contreras (2010) and Törrönen (2004) studies, either. Indeed, it is one of the most prevalent 'weaknesses' in the whole of CDA research. This characteristic is problematic in the field because it presents what Koller and Mautner (2004, p. 218, as cited in Baker et al., 2008) call a 'hidden danger' because, "...the reason why the texts concerned [in a given analysis] are singled out for analysis in the first place is that they are not typical, but in fact quite unusual instances which have aroused the analyst's attention." Because of its prevalence in the published research as well as the glut of critiques regarding it, the 'liberal' selection of texts is one of the first aspects of CDA research to have been effectively addressed by means of an alternate methodological approach.

2.3. Corpora and CDA

One factor which all of the studies mentioned above have in common is their basis in traditional CDA. That is, they look at different aspects of public discourse, but all share very specific methodological characteristics both in their gathering of texts and in their

analysis of those texts. Despite the fact that all CDA analyses are conducted using a corpus in the traditional sense (that is, a principled collection of texts), until recently there have been very few which have made use of electronic corpus analysis programs in their research. As was discussed above, the use of electronic corpus tools is more prevalent than ever. This is not only because of the sampling ease which these programs provide the researcher, but because electronic corpora provide a statistically based, verifiable, and—above all—consistent and objective research tool to investigators. Despite the advantages of using electronic corpus tools, the use of such programs has only recently become commonplace in CDA research.

In fact, the first prominent large-scale instance of pairing of CDA and corpus linguistics in a study was the Methodological Synergy approach, introduced by Baker et al. (2008). Although there have been other studies in the past which paved the way for the Methodological Synergy approach (see Baker et al., for discussion), this approach was unique in that it proposed the use of electronic corpus analysis programs as a means of combining corpus linguistics and CDA with the finality of demonstrating that the two approaches (CDA and Critical Linguistics) could be complimentary tools in carrying out research. Because this was one of the first attempts at marrying the two methods, the Baker et al. study—like the current study—was exploratory in nature. The study focused on the discourse present in printed news in England. The study was conducted using a 140 million-word corpus of news reports about British refugees and asylum seekers. Being as the study focused more on exploring a new approach than on the analysis itself, the main offering of the Baker et al. research was the conclusion that such studies are feasible and can potentially aid in the expansion of the field. The study did just that. As such, the most

important feature in the Baker et al. study was an overview of how such a methodology could be employed and how the results could be measured and analyzed by researchers.

Following the publication of the Baker et al. (2008) paper, one other author appears in the literature as having utilized their recommendations (at least under the same banner). Salama (2011) used the Methodological Synergy approach to analyze and interpret collocations in texts dealing with Wahhabi-Saudi Islam. Salama selected key words from two books regarding the subject and conducted a CDA-based collocation analysis in order to examine frequent (positive and negative) lexical uses within two corpora. Salama's research examined these concordances and used them to draw statistical conclusions about the uses of 'Wahhabi/Wahhab' within the texts examined. Based on the co-occurrence between selected node words and their collocates within the corpus, Salama was able to draw conclusions about LIs' use and significance within the texts; thus shedding light on underlying ideological motives within the corpora.

In Salama's (2011) study, the main focus of analysis was on the use of collocates in relation to paradigms. That is, the examination of frequent lexical occurrences as they relate/contribute to overriding views within a society. The study was carried out using two books regarding Wahhabi-Saudi Islam and was based on the analysis of the collocations of statistically prominent node words using Key Word in Context (KWIC) analysis. Despite these methodological measures, the research was still based on ideological interpretation in that both texts which were examined were selected due to their portrayal of a certain type of ideology (they were chosen due to the presence of antagonism in both texts). While Salama does compare language use in the corpora to a general corpus of American English (the *Corpus of Contemporary American English*) (a solution to many methodological problems

in CDA, proposed by Stubbs in 1997), the analysis is based on the antagonistic language used within the two books examined. As a result of this, the research findings highlight the very linguistic antagonism in both books which was acknowledged from the start of the research. In this respect, the Salama study serves as an exploration of the use of SP in corpus assisted CDA research. Also, the SP-based methodological tools Salama does use, are deployed in similar ways to many past studies. That is, node words were selected and were then searched with the intention of delineating their overall prosody within the corpora. Because of this, Salama's results—though immensely helpful in informing the methodology used in the present study—must be viewed critically in that the research effectively highlighted the antagonistic language which was already recognized in one way or another in beginning the research, and which was deemed a notable feature in the corpora from the beginning. Nonetheless, the Salama study is one of the most vital studies published to date in regard to the current project. While it can be criticized for focusing on antagonistic language within antagonistic corpora, it is one of the only prominent articles (at least at the time of this writing) which has taken the extra step—being added to the present study—of combining semantic theory (namely semantic prosody) to the corpus-based methodology set forth by Baker et al. (2008) with Methodological Synergy.

Regardless of the 'weaknesses' highlighted in the discussion of the above two studies (Salama, 2011; Baker et al., 2008), they represent two very important steps in achieving a more even-handed and objective methodology for conducting CDA research. The research carried out by Baker et al. serves to lay the groundwork for the methodological change employed here by showing that CDA can effectively be complemented by corpus linguistics tools; while Salama (2011) took this methodological pairing a step further by

adding the use of collocations to the analysis process, thus—in theory—making the analysis more objective. However, where both studies seem to come up short is in relation to the selection of the data used. Even after beginning to integrate corpus linguistics into CDA studies, the data analysis has remained largely subjective and ideologically influenced. That is, although data collection approaches have begun to change, the data interpretation has not. Additionally, while Salama's research is more methodologically even-handed than Baker et al., it relies on the discourse found in one very specific area of language use (the books analyzed); at the same time, Baker et al. looked at a broader language area but examined the collocates of very specific, pre-determined LIs. Although different researchers have looked at these characteristics of the field in different ways, some of the most promising advances have come as a result of the use of semantic theory. Indeed, while both studies offer very important ideas for carrying out more objective forms of CDA research, they each have certain methodological weaknesses (even though they are small). However, in the current study the strengths of each approach are used along with approaches based in semantic theory in order to propose a better way to carry out similar analyses.

2.4. Application of a Semantic Approach

As can be seen in the preceding discussion, CDA has traditionally focused on language from an overtly subjective perspective. This is true both from the perspective of data selection and from that of data analysis. While studies which have utilized corpus technology have served as a response to this 'problem,' the issue of analysis remains an important one in being able to establish a truly objective form of carrying out CDA research. This being so, despite the fact that it has not been very frequently used in

research, there is little evidence to say that semantic theory cannot be applied to a CDA-influenced methodology, as has been shown by Freake, Gentil & Sheyholislami (2011), Salama (2011), Baker et al. (2008) and Orpin (2005). Using these studies as a basis, the present research intends to pair SP (the advantages of which are discussed below) with elements of Methodological Synergy (Baker et al., 2008). The combination of these two approaches will allow for CDA-style research to be carried out with minimal input from the researcher—thus responding to one of the principal weaknesses in CDA methodologies.

One of the most prominent instances of a fusion between Methodological Synergy and theoretical linguistics was a 2010 study by Prentice which looked at Scottish nationalist discourses using semantic tagging. This study looked at the language used by Scottish nationalist sources in portraying Scottish independence movements. Although normally this sort of study would be carried out with a simple textual analysis, Prentice utilized an electronic corpus to semantically tag certain linguistic features within the texts. This allowed Prentice to focus not on the interpretation of a few scattered instances of overt nationalist language, but on the general use of certain features (the use of *Scot*, for instance) throughout an enormous array of texts, while simultaneously examining how these features were presented in order to ‘construct’ a certain perspective or frame. This approach was taken by Prentice, “...on the grounds of [CDA researchers’] subjective application of a coding system that has been created to fit the data under investigation,” (p. 408). Prentice goes on to suggest that,

this criticism is only partially addressed by applying one’s categorization system to a large volume of textual data, as there still exists a degree of subjectivity in the analyst’s application of the categorization system. Automated semantic tagging,

which offers us one possible way around this problem, and would allow relatively straightforward replication, involves the application of a computerized semantic coding system. (p. 408)

Semantic theory is one of the pillars of theoretical linguistics. Since semantic theory posits that LIs carry inherent meaning in any language, it follows that LIs are responsible for the communication of linguistic messages and, by proxy, the sorts of underlying discourses which are looked at in most CDA studies. Since semantics is generally seen as standing apart from context-based interpretations of language, semantic meaning is transparently present; its presence, then, dictates the underlying meaning present in a given use of language (either written or oral). To date, semantics has not been utilized extensively in studies dealing with CDA issues. Notwithstanding, there is a great deal of potential for its application in such areas. This is particularly true with regard to cross-linguistic studies.

Indeed, the present study is designed to experiment with and test this very application of semantic theory. Through the fusion of the approaches discussed in the preceding sections, it is proposed that a coherent, consistent, practical methodology can be created through which media discourse can be examined critically, but objectively.

One of the most obvious applications of semantic theory to CDA is through the use of semantic prosody. Generally speaking, semantic prosody in itself is what Stubbs (2007) calls, "...a relationship of habitual co-occurrence," (p. 1) or what Morley and Partington (2009) have described as 'evaluative harmony.' This is to say that—since semantic prosody is based on the analysis of collocations—it is predominantly focused on the idea that individual LIs have certain, consistent connotations (whether they be 'good,' 'bad' or

‘neutral’). Further, these LIs and their respective collocates are seen to both have an effect on and be affected by the linguistic environment in which they occur (Zhang, 2010; Morley & Partington, 2007; Stubbs, 2001b). Although the general understanding of so-called evaluative harmony is somewhat divided insofar as scope (Hunston, 2007), there is general agreement in the field that LIs and their respective collocates both have an effect on and are affected by the linguistic environment in which they occur (Zhang, 2010; Morley & Partington, 2007; Stubbs, 2001b). In this respect, semantic prosody is well-suited to CDA research—especially in relation to the cyclical relation between discourse and reality which is discussed above (Carvalho, 2008) and in identifying frames.

Despite this potential for application, the studies which have been carried out using a semantic prosody-based methodology have been relatively limited. Since Louw coined the term ‘semantic prosody’ in 1993, the vast majority of studies which have been carried out using a methodology featuring semantic prosody have been dedicated to the analysis of single words or single phrases. While this is not the case in every instance of semantic prosody research (see, e.g., Ping-fang & Jing-chun, 2009; Zhang, 2009; Partington, 2007; Xiao & McEnery, 2006), there are relatively few studies which do not focus on semantic prosody according to use and semantic environment (Yusuf, 2010; Stubbs, 2005) or metaphor (Fillmore et al., 1988; Oster, 2010). Despite the fact that—at least at first glance—these types of semantic prosody studies are unrelated to the application being proposed here, they serve an important purpose in that they lay out the potential strategies, approaches, and forms of analysis that can be used in carrying out a study which utilizes a semantic prosody as a methodological tool.

At its most basic level, SP is the analysis of collocations, what Fillmore et al. (1988)—in reference to idioms—call, “familiar pieces familiarly arranged,” (p. 510). That is, groups of words which consistently occur in groups within a given language. John Sinclair (1991)—one of the founding fathers of corpus and collocation research refers to two principles in collocation interpretation: the open-choice principle and the idiom principle. Both of these principles acknowledge that co-occurrence is a vital and universal feature of language; however, whereas the open-choice principle views language as a set of ‘slots’ waiting to be filled, the idiom principle conceptualizes it as a system in which certain combinations do not appear randomly (thus accounting for the idea of familiar arrangement put forth by Fillmore et al.).

To date, the most common usage of semantic prosody has been the analysis of idiomatic expressions and metaphors. While there have occasionally been innovative applications (Salama, 2011—in terms of the study’s use of SP as a complement to CDA), analyzing the prosody of individual words within large corpora to determine their most general prosodies remains the most accepted application. However, whereas the analysis of metaphoric language focuses on seemingly unchangeable units of language, the examination of the collocates of individual words searches for more ‘flexible’ patterns within language. The analysis of single words as they occur throughout hundreds of contexts is one of the most frequently used forms of semantic prosody investigation. Indeed, it is this type of study which gave birth to ‘Semantic Prosody’ as a theory; it was actually Sinclair’s realization in 1987 that LIs such as *happen* and *set in* almost always co-occurred with ‘negative’ events when looking at the Cobuild corpus which sparked the idea that semantic prosody might not only exist, but be systematically present within language

(see Zhang, 2010; Partington, 2004 for discussion). This, in turn led to Louw's famous 1993 study examining a few different LIs as well as coining the term 'semantic prosody' in reference to Firth's (1957) use of prosody to describe phonological characteristics which carried across words and sentences in speech giving the approach, "...its first definition, a 'consistent aura of meaning with which a form is imbued by its collocates'" (Louw, 1993 as cited in Zhang, 2010, p. 190). While Louw's study focused on the presence of irony in texts the field has since expanded to focus more prominently on questions of positive and negative connotations regarding habitual use of certain words in language.

These studies have confronted many uses of language, from looking at word uses as varied as the semantic prosody (SP) of the spoken use of *robot* (Yusuf, 2010) to things as simple as the use of the lexical item *fear* as a collocation used to portray emotions (Oster, 2010). However, regardless of the fact that there would seem to be a new crop of semantic prosody-related studies being published, the majority of readily available investigations continue to focus on broad, descriptive aims on the one hand (Stubbs, 2005; Partington, 2004) and research which shows more explicitly the presence of SP in language by using analyses focusing on irony (Partington, 2007) and cross-linguistic comparison on the other (Zhang, 2009; Xiao & McEnery, 2006).

Despite the fact that the field may seem broad in its approaches, it is actually quite limited in that the majority of studies simply focus on looking at semantic prosody and describing it within a given corpus. That is, there have been few practical applications using semantic prosody. Part of this is because the field is so young; however, this lack of application is also due to the sheer enormity of what can be looked at using SP.

Nonetheless, some practical investigations have been carried out and these studies—along

with descriptive accounts of semantic prosody—serve to lay a solid foundation on which to base practical applications of semantic prosody. This is principally due to the fact that there has not been a solid proposal made regarding how to go about undertaking a semantic prosody study. This has, as a result, led to a general lack of methodological consistency within published studies.

Stubbs (2005), for example, carried out a study of Joseph Conrad's book *Heart of Darkness* using methodology very similar to what is alluded to in much of the published literature regarding semantic prosody. While Stubbs' research focused on themes very similar to those found in discussions of semantic prosody, the term itself does not appear once in the article. In fact, it does not even appear when discussing the contributions which John Sinclair made to the field. Nonetheless, the article is a fine example of the potential for application which semantic prosody possesses. Stubbs uses corpus linguistics to dissect the language use employed by Conrad, eventually culminating in a linguistically based discussion of the literary style employed in *Heart of Darkness*. Although the study does not explicitly address semantic prosody, many of the findings are reminiscent of semantic prosody-based research (i.e. Conrad's use of antonymous pairs and lexical contrast).

Although the Stubbs study (2005) is likely one of the most practical studies on semantic prosody-related themes, there are other semantic prosody-specific studies which feature interesting applications of the approach. Yusuf (2010), for example set out to find the prosody of the node word *robot*. Using corpus analysis tools, Yusuf found that in the vast majority of cases in which *robot* is found in spoken language it possesses positive prosody characteristics stemming from the fact that robots are used to aid humans and facilitate work. While there obviously are not a great many studies which specifically focus on

semantic prosody as a practical methodology for linguistics research, the studies which have been published do serve to establish a base on which to construct a new methodological approach to CDA using SP methodologies.

With this in mind, the present study combines some of the methodological tools discussed here (in relation to SP) and above (in relation to CDA and corpus linguistics). Through this synthesis, the study is able to focus on relatively small, stratified random corpora from a specific language area and carry out a CDA study of language use within these corpora using SP in a novel way. The methodology which was used here and the justification for it are discussed below.