The Role of Citation in Interdisciplinary Discourse: an Investigation into Citation Practices in the Journal 'Global Environmental Change'

By

Fahad Saleh Aljabr

A Thesis Submitted to the University of Birmingham in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

School of English, Drama and

American & Canadian Studies

College of Arts and Law

The University of Birmingham

September 2018

Abstract

This thesis proposes an innovative model for citation analysis and applies it to 1186 citations derived from twenty papers from one interdisciplinary journal: Global Environmental Change. The main aim of this thesis is to build, not to quantify, a model which facilitates understanding of how citations act, and are acted upon, in citing texts. The model builds on, extends and modifies certain aspects of some existing models on citation form, stance and function. This thesis argues that stance and function are different but related concepts in the analysis of citation. They operate in different directions and, when combined, can reflect the role of citation in the citing text.

In order to achieve a fine-grained understanding of the role of citation, citations are analysed within and beyond the level of the statements in which they occur. To achieve this, a new level is proposed for the analysis of citation function: the 'citation block'. In this thesis, it is argued that citations operate in different directions within and beyond the proposition-level. The current thesis aligns and compares analyses at the clause- and block-levels for every citation. This alignment results in the identification of conventional and unconventional patterns of citing.

The model is applied to four sub-corpora of texts from two time periods and representing the more 'science-like' and 'social science-like' papers in the journal. The text-based analysis demonstrates the complexity of citation practices in interdisciplinary discourse. Overall it is suggested that in this journal the 'social science' papers over time have become more similar to the 'science' papers. The results also show variation in citation practices between the individual selected papers in each sub-corpus. This variation is attributed to the interdisciplinary nature of GEC. The proposed model has the potential to be used to investigate variation in citation practices beyond interdisciplinary discourse, within and between disciplines or genres.

Dedication

- *To my family; with all my love and deepest gratitude*
- *To my most loving and amazing wife who has stood by me through thick and thin*
- *To fatherhood; the best thing that has happened to me. To my little sons Saleh and Salman, with sincerest apologies for the dull weekends and for all the occasions when I had to go to my office and leave you both alone.*

Acknowledgements

All thanks and praise be to Allah, the Almighty for enabling me to pursue the degree of PhD and successfully complete this thesis.

The development of this thesis has involved many family members, academic guides and friends; whose support and guidance have been crucial throughout the many different stages of researching and writing this thesis.

First and foremost, it has been a great honour and privilege to have been supervised by Professor Susan Hunston OBE. I am, and will forever be indebted to Prof Hunston for her valuable support, constructive criticism of my work and suggestions on ways to improve my research. Every moment she spent reviewing and commenting on my work is deeply appreciated. Her guidance, detailed feedback and extremely helpful advice throughout the different stages of composing this thesis have been extremely enlightening. I would also like to thank her for giving me access to, and updates on the valuable resources originally compiled for the IDRD project (Thompson et al. 2017). This thesis would certainly not have been completed without her unwavering support and guidance throughout the last five years. I would also like to acknowledge with immense appreciation, the work of proofreading this thesis undertaken by Paul Daniel.

I would like to express my sincerest gratitude to the Government of Saudi Arabia, along with the University of Ha'il for awarding me the scholarship and giving me the opportunity to pursue my PhD.

Last but not least, I am profoundly indebted to my family. I am especially grateful to my parents and my brothers and sisters for their prayers and constant support throughout my PhD journey. My deepest gratitude must surely go to my wife and to my little heroes; my beloved sons, Saleh and Salman. I am at my wits' end as to how possibly I could even begin to thank my dear wife who has been unflinching in her support for me over these last few years. This arduous journey would not have been remotely possible without her fervent support, encouragement and most importantly, her patience throughout the many difficult times we have had to jointly contend with.

Table of Contents

List of Figures

List of Tables

Chapter One

Introduction

1.1 Overview: research background and context

This thesis proposes a model for citation analysis and applies it to a corpus drawn from the journal entitled Global Environmental Change (hereafter GEC). The main aim is to analyse citations for a better understanding of how they act, and are acted upon, in the citing texts. To achieve this aim, the current study proposes a model that builds on, extends and modifies certain aspects of some existing models (see Chapter 3). The model incorporates resources from within the form of citation (following Swales 1990), the writer's stance towards the cited authors[1] (following White 2003; Martin and White 2005 and Coffin 2009) and the citation function (using some terms as introduced by Harwood 2009).

Broadly speaking, linguistic research into citation practices focuses on the citation form, the writer's stance and/or the citation function. The first strand of research identifies different types of citations with respect to the citation form. From this particular viewpoint, citations are divided based on their syntactic realisation and are described as 'integral' or 'non-integral'. In integral citations, the author's name constitutes part of the cited proposition. On the other hand, in the non-integral citations, the author's name does not play any grammatical role in the construction of the statement in which the citation occurs. Such studies are form-based and build on the work of Swales (1990; see Chapter 2, Sub-section 2.3.2.2).

Another strand of research investigates citations in terms of the stance taken by the writer. Such studies build on 1) Thompson and Ye's (1991) work on evaluation in reporting, 2) the dialogic view of citation where citations are viewed as a form of engagement between the writer, cited authors and the anticipated readers, as particularised within the Appraisal theory in the systemic functional linguistics (SFL) tradition by White (2003) and Martin and White (2005) or 3) evaluation in attribution, as presented by Hunston (1989, 1993, 1995, 2000). See Chapter 2 (Sub-section 2.3.3) for a detailed discussion of the literature on stance in citation.

The third strand of citation research is concerned with citation functions. Research on citation function dates back to 1975 when Moravcsik and Murugesan first investigated

citations based on their role in text, as opposed to the writer's motivation for (non-)citing which has been extensively discussed in Library and Information Science (LIS) research (see Section 2.2). The linguistic research into citation functions could be divided into either text-based or interview-based; dependant primarily on the method utilised for eliciting the citation functions (Sub-section 2.3.4). What is noteworthy in this research is that citation function is often conflated with the writer's intention or motivation for (non-)citing (see Sub-section 2.3.4). In this thesis, following Petrić and Harwood (2013), I draw a distinction between the function of citation (the observed role of a citation) and the writer's motivation (reason for citing).

A few studies combine the analysis of citation form with tense and/or stance in reporting (e.g. Swales 1990: Thompson and Ye 1991; Hunston 1993, 1995; Pecorari 2013). Another strand of research adds a functional aspect to Swales' formal distinction (e.g. Thompson and Tribble 2001: Thompson 2001, 2005a, b). For example, Thompson and Tribble (2001) establish that the form of a citation could be indicative of its function. They identify seven functions of citations based on the form-function correlation (see Chapter 2, Sub-section 2.3.4).

There is a lack of research into the correlation between the writer's stance and citation function. Very few studies (e.g. Coffin 2009 and Ali 2016) investigate stance and function in citations. However, there is no alignment between stance and function categories in the aforementioned studies. This thesis starts with the assumption that such alignment would facilitate comparison of stance and function categories (see Chapter 4, Section 4.3 for examples) and more importantly, reflect how citations are evaluated and concomitantly employed in the citing texts.

The proposed model can therefore be seen as a triad system which aligns selected categories of form, stance and function of citations. In order to achieve the aim of reflecting on how citations act, and are acted upon, in texts, I analyse citations at two separate and distinct levels. Firstly, citation form and stance are analysed at the clause-level of the citing statement. In terms of citation function however, the current thesis introduces a new level for analysis, which is what I term the 'citation block'. Simply put, it is argued in this thesis that citations operate in blocks and for one to understand how they act in texts, it is insufficient to analyse their function at the clause-level. See

Chapter 3, Sub-section 3.5.2 for details related to the definition and identification of blocks.

To the best of my knowledge, there is no evidence of any previous research having undertaken the analysis of each citation at the clause- and block-levels. In this thesis, and in a previous pilot study, this form of analysis proved to be useful in helping to bring about a more accurate understanding of how citations operate in texts. Once the two levels have been devised, it was necessary to identify and align the form, stance and function categories in order to facilitate comparison between these categories.

There is unlikely to be divergence of views over the identification of the two forms of citation. Therefore, the formal distinction was adopted as the starting point for investigation. In order to investigate if the stance and function categories can be realised in both forms, I had to come to a carefully considered decision on the choice of stance and function categories. Beginning with stance, Coffin's (2009) compelling summary and extension of Martin and White's (2005) work on the engagement system was taken as the starting point for investigating stance in citation. Coffin introduces a list of four stance types used in citation, namely: 'endorse', 'contest', 'acknowledge' and 'distance'. These four stance types are taken in this thesis to represent the basic stance types in citation.

The next challenge was the identification of function categories which could be compared and aligned with the four stance types. Harwood (2009) lists eleven functions of citations. Four of these functions are closely related to the pre-identified stance types. These are: 'Support', 'Engage', 'Signpost' and 'Position'. It is important to note that whilst I adopt Harwood's terminology, I do not follow his methods for a number of reasons and these are enumerated in Chapter 2 (Sub-section 2.3.4) and Chapter 3.

Incorporating the aforementioned categories at different levels resulted in the identification of conventional and unconventional combinations of citations based on co-occurrences of form, stance and function categories (see Chapter 4; Section 4.3). It is also argued that such combination has different implications for pedagogy, citation theory and LIS and Corpus Linguistic research. These implications are discussed in Chapters 7 and 8.

At this stage, it could be argued that the multi-level analyses of citations in terms of form, stance and function have various implications for research into citation value.

This thesis argues that clause-level analysis of stance or even function in citations can in fact be deceptive and not representative of the citation role in the citing text. In the early stages of this research, Hunston (2015, personal communication) emphasised a publisher's (Elsevier UK) concern about the lack of research into citation value. This concern was in-line with personal interest in citations and their roles in texts and also shared, was the scepticism about citation counts (particularly h-indices) and their role of reflecting scientific contribution and academic impact. This scepticism is of course not unique to this research. In fact, a significant level of criticism has been levelled against quantitative measures of scientific contribution in LIS and linguistic research (see Chapter 2). Research in this thesis was initiated and conducted with an aim to find ways which can better reflect the citation role in text. Having said that, it is important to note that this thesis seeks to provide a richer understanding of how citations operate in texts. This, in turn, has various implications for citation theory, including research on citation value.

This thesis is deeply indebted to the Interdisciplinary Research Discourse (IDRD) project (see Thompson et al. 2017) for providing resources originally compiled for that project. The IDRD project compiled a corpus comprising 11,201 papers from eleven journals, the main journal being GEC (see Section 3.2 for further details on the IDRD project and 3.3 on data collection methods). Thompson et al. (2017) compiled 673 papers from GEC between 1990 and 2010 and divided them into six clusters, later called "text constellations" (ibid: 13). Papers in each constellation share sets of features and represent a distinctive approach to environment-related issues.

For this thesis, I have selected twenty papers, comprising 1186 citations, which represent the two most diverse constellations identified in the IDRD project (constellation 1 'Quantification and measuring' and constellation 5 'Personal voices'). Each of the 1186 citations was analysed at the clause- and block-levels and coded in terms of form, stance and function. The twenty selected papers were chosen from GEC in order to investigate citation practices in GEC as an interdisciplinary journal and to compare the results of applying the proposed model to those reported in the IDRD project (by Thompson et al. 2017).

The IDRD project did not investigate citation practices but the constellations were identified based on the occurrence of pre-identified linguistic features from Biber's

(1988) multi-dimensional analytical (MDA) model. It was expected, given what was known about the interim results of the IDRD project in 2015-2016, to find variation in the citation practices between the two constellations. The results are discussed in light of the IDRD project in Chapters 6 and 7.

In this thesis, the corpus was further divided into four sub-corpora, each representing one constellation (either 1 or 5) and one time period (either 1990-95 or 2008-10). The goal was to identify diachronic and synchronic variation between the selected papers and more importantly to reflect on citations and their roles in the texts (see Sub-section 3.3). Chapter 5 furnishes the quantitative results of comparing the constellations and the relevant time periods. It is crucial to note that the aim is not to make generalisations. Rather, the main aim is to show tendencies in the selected papers and to give an indication of diachronic convergence and divergence within each sub-corpus. The occurrence of form, stance and function categories is compared across the four sub-corpora in Chapter 5 and discussed in Chapter 7.

1.2 Rationale for the study

In this section, I shall justify my decision to study citations in the journal GEC. I shall also provide a brief example to demonstrate the importance of the block as a new analytical unit. I conclude this section with a justification of the choice to study citations in research articles as opposed to other genres in academic writing such as textbooks or student writing at postgraduate level.

The current thesis addresses what is clearly an important issue in academic discourse: the nature and practice of citation. In the literature, there has been extensive research into the form, stance and/or function of citing. Citation practices have been compared in various contexts (e.g. across disciplines, as in Thompson 2001, 2005a who analyses citation practices at PhD level in Agricultural Botany and Agricultural and Food Economics; or between different language groups, as in Ali's 2016 contrastive analysis of citation practices in Sudanese and British PhD theses in Applied Linguistics; or even between different genres, as in Samraj's 2013 analysis of citation practices in discussion sections of research articles [RAs] and Master's theses). However, there is a lack of research into citation practices in interdisciplinary discourse. This study aims to address this knowledge gap and thus investigates citation practices in GEC as a well-established interdisciplinary journal (as identified by the journal's publisher Elsevier UK, cited in

Thompson et al. 2017) which targets an interdisciplinary audience (Hunston 2018, personal communication).

The term 'interdisciplinary' is problematic as it has different meanings in the literature. With respect to academic journals, for example, it is often used to refer to joint disciplinary journals which receive contributions from two or more disciplines (e.g. journals of Computational Linguistics, which draw on computer science and linguistics). There is a lack of research comparing citation practices in different disciplinary origins on the one hand with the joint discipline they produce on the other (though see Muguiro's 2015 analysis of citations in Educational Neurolinguistics). Such research would facilitate understanding of how each discipline has contributed to knowledge construction in the joint discipline.

When it comes to GEC, the difficulty for the researcher is that it is not always apparent which disciplines individual papers belong to. In other words, GEC receives contributions from a variety of disciplines. However, since GEC does not identify the disciplinary origins of the various contributions, it is not easy or even possible to assign papers to specific disciplines.

During the early stages of this research, the initial plan was to compare citation practices in a number of papers from different disciplines in GEC. However, due to the aforementioned reasons, it was not possible to study or compare the different disciplinary norms in the chosen journal. The immediate challenge was then to find a way to select papers from the interdisciplinary journal GEC that reflects its interdisciplinary nature. In a pilot study of some GEC papers (2015), I noticed that some papers are science-like, whereas others are more social science-like. This observation was later confirmed by Hunston (2016, personal communication) as a member of the research team in the IDRD project, who demonstrated that some of the papers in GEC represent the scientific component of the journal whereas others represent the social scientific component (the findings of the IDRD project were published in 2017 by Thompson and colleagues, including Hunston). GEC specifically notes on its website (retrieved 5 May 2018) that it welcomes papers with a social scientific component.

It was thus decided to select papers which represent the scientific and social scientific components of the IDRD corpus. The authors of the project (Thompson et al. 2017)

identified 53 linguistic features from Biber's (1988) multi-dimensional model and then assigned each of the 673 papers in GEC to one of six constellations based on its dimensional scoring of the pre-identified linguistic features. From the six constellations, two constellations showed the greatest variation in their dimensional scores: these are constellation 1 (Quantification and measuring) and constellation 5 (Personal voices). The former constellation contains papers which represent the scientific component of the corpus and the other constellation represents the social scientific component (Thompson et al. 2017).

Although the IDRD project did not investigate citation practices in the constellations, the observed variation in the project led to the expectation in this thesis that the two constellations will show variation in terms of the use of citations. In order to investigate this expected variation in citation practices, I decided to sample the two most distinct constellations in the IDRD project to compare the results of this thesis to those reported in the IDRD project.

In summation, this thesis uses two constellations from the IDRD project (that is, constellations 1 and 5) as proxies for disciplines even though they are not actually the same as disciplines nor were they originally identified in the IDRD project to reflect disciplinary variation in GEC. Constellation 1 papers are taken to represent scientific papers and constellation 5 papers are taken to represent the social scientific papers. In other words, I use the results of a corpus study of interdisciplinary discourse to make a principled collection of texts. The selected papers are then textually analysed in order to apply and to revise the proposed model. The selected papers together form the corpus of the current study.[2]

Due to the in-depth nature of analysis in this thesis, it was necessary to limit the number of papers chosen for investigation in order to facilitate text-based analysis of citations-in-contexts (see Adel 2014 for a discussion on how to select a subset of quantitative studies on which to apply qualitative research). The initial plan was to include as many papers as possible from the two constellations. However, I had no control over the distribution of papers in the constellations. The papers were not evenly distributed across the constellations as some constellations contain a considerably smaller number of papers than others (e.g. constellation 1 contains 118 papers and constellation 5 contains 35 papers, only 5 of which were published between 1990 and 1995). As a

result, I had to make a decision on sampling the two constellations in a way which facilitates identifying diachronic and synchronic convergence or divergence in the selected papers. See Chapters 3 and 5 for further detail on sampling and synchronic and diachronic variation.

Once the papers were selected, the main challenge was the identification of the context for analysing and comparing the various aspects related to citation analysis (e.g. form, stance and function). It is notable in this regard that most citation classification schemes trace citation functions at the clause-level of the statements in which citations occur. This study adopts a different approach and proposes a new analytical unit, particularly for the investigation of citation functions: the citation block. As noted above, the main argument is that it is insufficient and often misleading to analyse citation functions at the clause-level, as previous research tends to do (e.g. Petrić 2007; Petrić and Harwood 2013; Dontcheva-Navratilova 2016). This view receives support from Sinclair (1985, 1986b; cited in Jabbour 1997: 141) who demonstrates that a writer's position on some occasions becomes clear as the text unfolds. He specifically writes:

> In writing it is commonplace for a writer to state a position that is contrary to the position he wishes the reader to adopt; the eventual effect can only be described when the artefact has unfolded in full (Sinclair 1985: 16; also in Sinclair 2004: 70).

This would suggest that a schema classifying citation function(s) should take account of an extended context. As noted in the previous section, the main argument in this thesis is that citations operate in blocks. The combination of clause-level analysis of citation form and stance together with block-level analysis of citation functions reveals the position(s) taken by the writer towards the cited information. Simply put, the writer's stance beyond the clause-level of the citing statement is taken as an indication of the citation function; how it operates in the wider context (see Sub-sections 4.3.1 and 4.3.2 for examples). Chapter 6 provides examples in which the citing writers simply acknowledge other sources at the clause-level of the citing statements. The same writers are often involved in critical dialogue with the cited authors or their material at block-level. In such examples, the writers Engage with the cited material.[3] Function analysis in this sense complements stance analysis and together they reflect how citations are used in a context wider than that of the clause.

To demonstrate the importance of the block, consider the following extract from Sinclair (1987: 3), represented here as Example 1.1 (the statements are numbered for ease of reference).

Example 1.1:

> [7.1] The slogan may have inspired the soldiers, but it didn't help Mao much. [7.2] Just six weeks after the Tangshan cataclysm, the Great Helsman died at 82. [7.3] To superstitious Chinese, his death so soon after the the [*sic*] earthquake was no surprise. [7.4] Traditionally, natural calamities were believed to signal the disfavour of Heaven with the ruling dynasty. [7.5] Floods, earthquakes and other portents were supposed to precede the demise of the emperor.
>
> [8.1] To more practical minds, the Tangshan disaster had a different meaning. [8.2] It showed that despite the imaginative efforts of Chinese seismologists, the science of earthquake prediction was still very much a hit-and-miss affair. [8.3] Since 1976, the Chinese have therefore redoubled their efforts to understand the titanic forces beneath the earth's surface. [8.4] Those studies, they hope, will one day enable them to forecast quakes with greater accuracy – and prepare for the next Tangshan.

Sinclair analyses this example as part of a wider text about an earthquake which hit Tangshan in 1976 and demolished the city. These two paragraphs (paragraphs 7 and 8) appear within a passage consisting of eight paragraphs in Sinclair's study. Sinclair is specifically interested in the interplay between averral and attribution in this example (see Chapter 2, Sub-section 2.3.1 for further discussion of averral and attribution). Here, I use these two paragraphs to give an indication of the significance of the block in revealing how the attributed information is evaluated and more importantly, how it functions with respect to the surrounding context.

In paragraph 7, the last three statements (7.3, 7.4 and 7.5) are attributed to 'superstitious Chinese'. The writer clearly distances themselves from the attributed information (*To superstitious Chinese, were believed to* and *were supposed to,* respectively). In other words, the writer does not share responsibility for the propositions 7.3, 7.4 and 7.5 and so the above statement (in 7.3) is an instance of distanced attribution. The use of the structure *To superstitious Chinese* clearly transfers responsibility for the proposition to another group. It is important to note that the writer's contradiction of the source's view in 7.3 is deferred to the next paragraph.

The attributed views in 7.3, 7.4 and 7.5 (to superstitious and traditional Chinese) stand in contrast with the writer's authorial position but this does not become clear unless we

expand the context to include the next paragraph (paragraph 8). Both 8.1 and 8.2 are attributed, but to *more practical minds*. The use of this comparative form and the word 'practical' indicates that the writer regards this interpretation of the disaster as more reliable than that of superstitious Chinese (in 7.3). Further, the choice of the reporting verb *showed* in 8.2 elaborates on the interpretation of the *more practical minds* and presents their view as "a generally accepted position" with which the writer is not expected to disagree (Sinclair 1987: 10). Likewise, I argue that the writer achieves Engagement with the attributed information in 7.3 by presenting another more reliable 'practical' position in 8.1. In other words, the writer aligns themselves with the second interpretation (in 8.1 and 8.2) via the use of the reporting verb *showed*. This reporting verb indicates the writer's *endorsement* of the interpretation in 8.1 and 8.2.[4] It could thus be argued that the writer reclaims or shares responsibility for the attributed information in 8.1 and 8.2. The same writer is involved in critical dialogue with the views posed in 7.3, 7.4 and 7.5.

To summarise, while Sinclair analyses the above example to discuss the interplay between averral and attribution beyond the sentence-level, I have reconsidered it to show the importance of the block as a new unit for analysing citation function. The aim of the current thesis is thus to understand how citing writers make alliances and position themselves towards cited arguments and, at the same time, how they employ the cited materials in the ongoing text at block-level. Combining the clause- and block-levels reflects how writers manipulate the use and evaluation of attributed information and can shed light on the textual purpose. The importance of the block becomes clear towards the end of this thesis. At this stage, I would argue that it is crucial for a better understanding of citation functions.

The identification of citation blocks has various pedagogical implications. In fact, the identification of blocks was inspired by personal interest developed during the course of employment as an EFL instructor at university level, and observation of an evident challenge faced by non-native university students regarding citing and presenting external voices both at the clause-level and beyond. In general, EFL students and novice researchers struggle with this crucial feature of academic writing (Harwood 2010). One of the aims is thus to investigate citation practices in published writing with the assumption that it will have various implications for pedagogy, specifically in the EFL context. The proposed model identifies various citation patterns (see Chapter 4) which

can be useful in sensitising students to successful citation practices and enabling them to reflect on their own practices in comparison with those of published writers (some of the implications are discussed in Chapter 7).

In order to investigate published writing, this thesis specifically targets citation practices in GEC RAs. RAs, as a genre of academic writing, have received more attention than any other genre in the literature on citation studies. This is due to the fact that RAs play a vital role in shaping and developing the ongoing debate in disciplinary knowledge. It could be argued that RAs have a wider global reach than other genres or forms of written academic knowledge, such as textbooks or student writing at postgraduate level. Large numbers of RAs are exchanged and compared every day because knowledge is constantly changing at a rapid pace. As such, the rate of published RAs reflects the pace of knowledge growth. Further, it is easier to trace how academic writers make alliances and judgments about other authors' claims and research in journal articles than in other forms of academic writing. This is partly due to their availability in electronic form and their length.

GEC is regarded as a successful interdisciplinary journal which targets a wide interdisciplinary audience. The citation practices were therefore expected to reflect the interdisciplinary nature of the journal and at the same time to be accessible to audiences from outside the field of environmental science, including the analyst with a linguistic background (and possibly students in future applications of the model).

1.3 Research aims and questions

As stated at the beginning of this chapter, a fundamental aim of this thesis is to achieve a better understanding of the ways in which citations are employed and reflected upon in the citing texts. A secondary aim is to analyse citations at different levels, mainly because a one-level analysis is insufficient to display the writer's evaluation and the citation function. Combining levels facilitates the understanding of citation roles in texts. The commonalities and variations between the results of each level can be examined, and this examination has in fact resulted in the identification of conventional and unconventional patterns (Chapter 4, Section 4.3).

In terms of the selected corpus, the current thesis aims to investigate the citation practices in some papers from GEC and to give an indication of synchronic and

diachronic variation between the selected papers which represent the two most diverse constellations, as demonstrated by Thompson et al. (2017).

The literature shows that students show a lack of understanding of why and how to cite (e.g. Harwood 2010). Therefore, it is hoped that the multi-level analyses will provide a model for language teachers and learners to deconstruct texts and to expose students to various ways of presenting external voices within and beyond the level of the clause. This exposure should enable students to see what effect their linguistic choices at the two levels (at the two levels) have on the reader.

The current study thus seeks to answer the following research questions in relation to the proposed model in this thesis:

1- Can the proposed model robustly reflect the role of citation in academic discourse?
2- What can the proposed model reveal about citation practices in the journal GEC and the synchronic and diachronic variation between the selected papers?

The first question is answered in Chapters 3, 4 and 6. The second question is answered in Chapters 5, 6 and 7.

1.4 Outline of the thesis

This thesis is divided into eight chapters. Chapter 2 provides a theoretical background to the thesis topic. It surveys the relevant literature from two perspectives: the first is related to information scientists and sociologists of science, and the second explores linguistic approaches to the analysis of citations.

Chapter 3 covers issues of data collection and sampling. It also introduces the proposed model and the different levels of analysis, and presents other aspects relevant to the analysis such as the identification of the block and the 'cited-cited relation' (i.e. the links which the writer draws between two or more external sources). Next, Chapter 4 exemplifies the model, by providing qualitative examples of the categories at each level. Due to restrictions in space, it is not possible to exemplify every pattern especially where analysis is conducted at the block-level. The combination of some stance and function categories is exemplified in this chapter and the resulting patterns are highlighted and defined.

Chapter 5 provides the quantitative results of the application of the various model categories. Due to the limited number of papers analysed, the results are not common to all GEC papers. Instead, they reflect tendencies in the selected papers. The focus was on devising a qualitative model that enables the understanding of citation relations within and beyond the citing text. This chapter highlights the diachronic and synchronic variations in the selected corpus. The results are then discussed in light of the findings from the IDRD project in Chapter 7 (see Thompson et al. 2017 in Section 3.2 for further details on the IDRD project).

Chapter 6 presents one exceptional paper that led to different outcomes in almost every aspect of analysis. This paper is therefore case-studied in Chapter 6 and a detailed analysis of its citations is presented as well. In the same chapter, the paper is compared to one typical paper from the same corpus to demonstrate how and why they exhibit this variation.

Chapter 7 discusses the findings of the study in terms of its various implications. These include implications for pedagogy, for citation theory including the relationship between stance and function and the applicability of the notions of dialogic expansion and contraction to citation functions. The findings are also discussed in terms of possible implications for corpus linguistics and information science. Finally, Chapter 8 summarises the whole thesis, highlighting the main contribution, arguments and findings from each chapter.

Notes

1- Adopting Thompson and Ye's (1991) terminology, I refer to the citer as the 'writer' and the citee as the 'author'.
2- Corpus in this sense does not refer to a large collection of electronic texts. It simply refers to the selected papers. Therefore, the results of applying the model are intended to reflect tendencies in the selected papers only.
3- In this thesis, Engage (with capital *E*) is used as a function of citation and is derived from Harwood's (2009) model to mean that the writer is involved in critical dialogue with the source. It is the extreme opposite of 'Support' as a function, see Chapter 3 for definitions and examples.
4- From now on, to distinguish stance and function categories, I shall capitalise the function categories and italicise the stance categories.

Chapter Two

Theoretical Orientation

2.1 Introduction

As briefly noted in Chapter 1, the current body of knowledge of citation practices is drawn mainly from two very different disciplines: Library and Information Science (LIS) and linguistics. In the former discipline, the focus of research has been mainly on improving retrieval performance and assessing the scientific contribution of individuals via citation indices (Bornmann and Daniel 2008). The focus therefore, has not been on language or knowledge construction, but on using citations as a gateway to understanding a larger phenomenon relating to research impact and citation networks.

However, the case is different for linguistic research into citations as linguists perceive citation as a linguistic phenomenon that may be analysed to shed light on how citers interact with other sources. Linguistic research into citation is largely motivated by the need to better understand citation practices and a concern for the rhetorical function of citation (e.g. Petrić 2007; Coffin 2009; Harwood 2009; see Section 2.3).

The purpose of this chapter is to trace the development of citation research in these two disciplines and outline how it has informed the current study. The next section (2.2) summarises the origin and advancement of LIS research on citation. The question that arises in this section is whether LIS research mimics in anyway the linguistic research that I will report in Section 2.3. In Section 2.4, I summarise the arguments of these two fields and explain how they contribute to the proposed framework in the current thesis.

2.2 Citation in Library and Information Science (LIS): a historical account

Citation is part of a more general academic practice of engagement with the research community. In linguistics, the analysis of citation has been mainly focused on the nature and purpose of citing. Applied linguists are familiar with studies of citations as within the broader tradition of investigating academic texts, but citation has also been studied from other perspectives. In addition to LIS, disciplines of Science and Technology Studies (STS) and Sociology of Science (SOS) have also made various contributions to the study of citation, as an indicator of research impact.

In LIS, a substantial body of the literature has addressed different aspects related to the study of research impact and consequently, various quantitative measures have been devised. These efforts date back to 1927 when Gross and Gross (1927) first used citation counts to evaluate the importance of scientific work. Since then, citation counts have been studied as a measure of journal development (e.g. Garfield 1972; Nicolaisen 2002), discipline development (e.g. Oppenheim 1995, 1997) and the scientific contribution of individual researchers or departments (e.g. Hagstrom 1971; Mahlck and Persson 2000). The use of citation counts as a means of measuring the scientific impact of individual researchers may be traced back to the 1960s (Ding et al. 2014 and Wan and Liu 2014). Thereafter, a number of different impact indicators have been proposed, the most popular being the h-index and its numerous variants (Wan and Liu 2014). Broad reviews of citation research in LIS and its development include Smith (1981), Small (1982), Cronin (1984), Pierce (1987), Baird and Oppenheim (1994), Liu (1993a), McCain and Turner (1989), Case and Higgins (2000), Bornmann and Daniel (2008), Ding et al. (2014) and Wan and Liu (2014).

Within scientometrics and from the perspective of STS, citations have been largely treated as indicators of reward in the scientific system (Luukkonen 1990, 1997). In other words, the number of citations of a given work was taken as an indication of its impact (the degree to which it has influenced scientific research). Leydesdorff (1998) notes that this impact enabled scientists to translate the cognitive aspect of citing to the system of social rewards in science. Some scientometric instruments that have been proposed to measure and assess the citing and cited correlation include 'co-words' (Callon et al. 1982), 'citation networks' (Mahlck and Persson 2000; Sandstrom et al. 2005) and 'co-word analyses' (Callon et al. 1982; Van Raan 1988). Each of the above indicators reflects methodological differences. However, the differences are irrelevant to the current study (for detailed reviews, see Cronin 1984; Liu 1993a; Leydesdorff 1998 and Bornmann and Daniel 2008).

The quantitative measures of scientific impact are generally based on counting the number of times that any given piece of research is cited in succeeding texts. Additionally, a number of studies in LIS have differentiated between the number of citations that one text receives and the number of mentions of the same source in one text (e.g. Hooten 1991). These studies propose that the number of mentions is highly predictive of the impact of the cited text (e.g. Pride and Knoth 2017a). However, such

studies do not provide qualitative evidence for their claims, and their claims stand in sharp contrast to the argument that citations are not equal (e.g. Voos and Dagaev 1976; Herlach 1978; Cano 1989; Zhu et al. 2015; Bertin et al. 2016). It is important to note that quantitative measures of frequency and mention of citations are incapable of distinguishing positive, negative and perfunctory citations.

Despite that, citation counting to date continues to receive attention as a measure of scientific impact, and various studies have argued for the importance of citation indices in assessing the cited work's contribution and influence on the scientific community (e.g. Van Raan et al. 2003). However, this view of research impact has faced prolonged criticism. In general, researchers interested in citation counts and their role in measuring scientific impact constitute two schools of thought. According to one school (e.g. Cole 2000; Van Raan 2004), "evaluative bibliometric analyses" (Bornmann and Daniel 2008: 46) suffice for assessing scientific impact. In other words, this group advocates that citation counts are consistent with other assessment measures, such as departmental prestige (e.g. Hagstrom 1971; Anderson et al. 1978), awards and Nobel laureateships (e.g. Cole and Cole 1967) and academic rank (for a complete account see Cronin 2005; Bornmann and Daniel 2008). On the positive side, citation counts are useful when analysing large amounts of data, time-efficient and relatively indicative of the impact of a piece of research.

In contrast, the other school of scientists casts doubt on the assumption that citation counts can successfully reflect scientific impact. This long-lasting criticism dates back to the 1960s (e.g. Garfield 1962, 1965, 1972, 1998; Woolgar 1991; Burrell 2003; Rabow 2005; Bertin et al. 2016; Pride and Knoth 2017a, b), and opposes h-indices and quantitative measures by demonstrating that they are insufficient to reflect scientific impact on their own. While they may be useful in measuring an individual's research influence and reach, they fall short of appropriately explaining the content and context of citations and thus the different contributions made by different citations (Lawani and Bayer 1983; Beed and Beed 1996). Instead, it is proposed that such quantitative measures should constitute only one, among many other, variables that can influence decisions regarding scientific impact or research value. One of the difficulties is that the probability of being cited relies on many factors such as time, field, journal, article, author/reader-availability and technically-related factors rather than on the quality of the research itself (Bornmann and Daniel 2008).

The criticism of citation counts and their role in assessing scientific impact has its roots in SOS (e.g. Merton 1973; Moravcsik and Murugesan 1975; Cozzens 1989), information science (e.g. Gilbert 1977; Shadish et al. 1995; Cronin 1982; Cano 1989; Amsterdam and Leydesdorff 1989; Maričić et al. 1998; White 2004; Zhao et al. 2012; Belter 2014; Zhao and Strotman 2014; Pride and Knoth 2017b) and more importantly in linguistics (e.g. Swales 1986, 1990). Cronin (1982) contends that the norms of citation are not precisely formulated which leads to uncertainty of what is being measured. Another important criticism is that citation counting does not take into account the very different nature of citations and the fact that each citation performs different role(s) depending on the context in which it occurs. As each citation is context-dependent and whilst context studies of citations have proposed different contexts and units of analysis (e.g. Brooks 1985, 1986; Bertin et al. 2016), they have not properly articulated a theory for approaching citations to investigate their observable roles in contexts.

This ongoing criticism has resulted in an increasing awareness and recognition of the need to improve such measures. As a result, the focus has gradually shifted from the study of citations as a means of improving retrieval performance to the study of citing behaviour. A substantial body of the literature has been geared towards the qualitative study of citer behaviour and the reasons for (non-)citing. The literature on citing behaviour (within LIS and bibliometrically inclined SOS) can be divided into three main areas, these are: 1) context studies, 2) content studies and 3) studies of citer motivation. These streams of research reflect methodological variation and are covered in great detail in Brooks (1988), Cronin (1982), Liu (1993a), White (2004) and Bornmann and Daniel (2008). Empirical studies on citer behaviour and motivation differ in terms of their methodological approach. They are generally either context- or content-dependent – that is, based on analysis of text – or 'emic', which is based on surveys and interviews with citers (Bornmann and Daniel 2008).

The first stream of research into citing behaviour, context studies, involves text-based analysis of the citation context. The context is generally defined as the "co-occurring text" around the citation (White 2004: 100). There is a tendency in such context studies in LIS not to explain the steps involved in deriving and coding citations in terms of the proposed categories (though see Teufel et al. 2006 on the identification of "cue phrases" (ibid: 107) such as *Further to the work of* and *We employed* in order to automatically

classify citation functions). In general, context studies investigate the "inter-document relationship" (Bornman and Daniel 2008:50-1). Such studies produced different classification schemes because they identified different codes and contexts. The limitation of such schemes is that they occasionally require expert knowledge to retrieve implicit meanings.

The second stream of research (which involves content studies) is a response to the limitations of context studies. In content studies (also known as "content analyses of citation contexts", White 2004: 100), there is no role for the analyst's expertise in coding and formulating the classification schemes. Instead, the "semantic content of the citing papers" (Liu 1993a: 379) is used to characterise the cited works. The analyst here relies on explicit textual signals in the citation context to classify citations. Examples of such studies include Hooten (1991) and Hanney et al. (2003). In LIS, this stream of research is important since the citation content can be analysed to indicate topicality, which is extremely useful for document retrieval. However, White (2004) shows that a common feature in all classification schemes is that they are "idiosyncratic and hard to code" (p. 100) which makes them difficult to replicate across different disciplines or genres. The above two streams of research reveal that the evaluations implicitly carried by citation counts are distorted because they do not distinguish *perfunctory* from *essential* citations (ibid).

The third stream of research is related to the motivation or reasons for citing. The pioneers in LIS research on citer motivation are Garfield (1962, 1965), Brooks (1985, 1986), Latour (1987), Vinkler (1987), Bonzi and Snyder (1991), White and Wang (1997, continued in Wang and White 1999) and Case and Higgins (2000). Because the motivations or intentions of citations are not always available to content analysts (Gilbert 1977; Bornmann and Daniel 2008), this research is characterised by its use of interviews (e.g. Collins 1999; Case and Higgins 2000) and surveys (e.g. Shadish et al. 1995) to elicit author motivations (some of the limitations of this method are given in Section 2.3.4). Despite recurrent attempts to classify citations based on their motives, there is an agreement that a proper comprehensive theory or classification scheme has not yet been formulated (e.g. Bertin et al. 2016).

Taking an exhaustive view reveals that the above classification schemes operate in isolation of each other and even when some researchers rely on certain aspects of

previous schemes (e.g. Chubin and Moitra 1975 being inspired by Moravcsik and Murugesan 1975), they tend to use the same methods and operate at the same level, in the very limited context of the citing statement. Remarkably also, the vast majority of motivation studies use the terms 'functions' and 'motives' interchangeably although there are obvious differences between them. Only a few linguistic studies (e.g. Petrić and Harwood 2013) make the distinction between function of a citation (the rhetorical role of a citation) and its motive (the reason for citing).

The emergence of citation indexing in the 1920s changed the practice of citation analysis (Leydesdorff 1998). In general, cited texts were taken as a lead to trace the history of a knowledge claim; citation indexing was introduced to shift the focus to citing texts, which were seen as providing opportunities to cite previous works. In this view, older texts compete for citations and citation counts were proposed in response to this, as a new measure of scientific impact. The focus thus shifted from the cited to the citing texts.

While citation indexing models were influential in changing the practice of citation analysis, they have many limitations. The main limitation, as noted above, is that citation counts oversimplify the different contributions made by different citations. Collins (1985) points out that citation counts do not allow inference regarding the impact of an article; the quality or lack of quality of a researcher cannot be measured from citation counts. Similar criticisms of citation counts and their validity as a measure of scientific growth and quality have been raised since the 1960s and 1970s (e.g. Cole and Cole 1967; Chubin and Moitra 1975; Moravcsik 1973, and Moravcsik and Murugesan 1975).

As a result of this criticism of quantitative measures of scientific impact, a parallel strand of research in LIS carried out qualitative analyses of citations (e.g. Lipetz 1965; Frost 1979; and Duncan et al. 1981). This consideration of the qualitative aspects over the years has significantly contributed to the field of citation analysis (White 2004). Quantitative measures need qualitative theoretical support in order to account for citing behaviour, and more importantly, citation value (for extensive reviews of LIS literature on citation analysis see Smith 1981; Small 1982; Cronin 1984; Pierce 1987; Baird and Oppenheim 1994; Liu 1993a; McCain and Turner 1989; Case and Higgins 2000).

From the 1970s onwards, there has been an increasing awareness of the social, cognitive and textual dimensions of citations (e.g. Leydesdorff 1998; Bornmann and Daniel 2008). However, research in STS and neighbouring disciplines with respect to citation analysis has not been a cumulative endeavour (Bornmann and Daniel 2008). Studies were conducted in isolation and each researcher, or research group, has developed its own perspective regarding the citation analysis query. This variation in design and purpose resulted in a variety of results and conclusions. The literature reports similar criticisms of citation classification schemes that are also attributed to the variation in design and purpose (e.g. Bornmann and Daniel 2008).

In an attempt to eliminate the need for manual coding, some context analysts opted for automatic identification and classification of citation functions. Most of this literature was geared towards understanding citation functions based on automatic extraction and classification of citations, a first attempt of which was proposed by O'Connor (1980) and followed by Hooten (1991), Maričić et al. (1998), Garzone and Mercer (2000), Hanney et al. (2003), Mercer and Di Marco (2004), Teufel et al. (2006) and and Bertin et al. (2016). Some of these studies are revisited in the review of citation function research in Section 2.3.4.

Garzone and Mercer (2000) attempted to automatically classify citations according to their functions (cf. Mercer Di Marco 2004). They embarked on a course of inputting complete journal articles in electronic form into a citation classifier which they had proposed in an earlier study (1996), and received a list of citations and functions as the output. In an empirical study, a test of the newly developed classifier on authentic data revealed that 80% of functions were accurately identified. Similar results had been obtained by O'Connor (1982).

While Garzone and Mercer (2000) acknowledge the "crucial" role of the context surrounding a citation (p. 339), their automatic extraction and analysis process is largely based on an analysis of citing statements. Furthermore, it assigns each citation with only one function, thus excluding those citations that are multi-functional (as also found in Harwood 2009; Samraj 2013; Petrić and Harwood 2013, and also shown in this thesis in Chapters 4, 5 and 6). Overall, Garzone and Mercer (2000) admit that their model could benefit from adding more "obvious linguistic elements" (p. 342) to facilitate the identification and coding of citation functions.

In response to the weaknesses in existing LIS coding schemes, Zhang et al. (2013) propose a Citation Content Analysis (CCA) codebook which aims to complement existing schemes. The CCA codebook is two-dimensional (citing and cited) and bi-modular (syntactic and semantic); in other words, it attempts to fill the gap in previous approaches by combining the two dimensions. The CCA codebook is based on a number of assumptions. First, that citations are unequal and thus they have different levels of importance or relevance. Moreover, context plays a vital role in the study of citations and thus bibliometric indices are insufficient. The second assumption is that qualitative analysis is inseparable from quantitative measures (e.g. citation counts) and therefore should be incorporated into them. CCA represents an endeavour to describe the "contextual relationship" (ibid: 1491) between the citing and cited works (and vice versa). Accordingly, CCA explores the role of individual and collective norms in citing behaviour.

The CCA framework targets three features of citations. The first is the numerical nature of citations, and assumes that research impact is measured in terms of citation counts, as in traditional approaches and h-indices. The second feature is literal, revolving around the notation that citations consist of words and these words can be studied implicitly or explicitly to unravel the citer's attitude, intellectual process and sentiment. The third feature is sociocultural, whereby the process of citing is a complex social system in which both individual factors (e.g. attributes and style) and collective norms (e.g. domain conventions) affect the way an author cites. Also, the motivations and intentions behind citing vary and are context-dependent, thus making this feature implicit and abstract. The authors are unsuccessful, I believe, in explaining and exemplifying the second and third features, because they do not take advantage of linguistic research (e.g. White 2003; Petrić 2007; Harwood 2009; Coffin 2009 and many others; see the next section). In general, their work seems to be more oriented towards information scientists.

A recent study on CCA, published in the same journal as the original paper by Zhang et al. (2013), reviews and reflects upon previous research on citation analysis, starting with traditional citation analysis of small sets of academic articles and progressing to consider modern computer-assisted semiautomatic analytical tools and techniques (Ding et al. 2014). The authors discuss current approaches to CCA and future potential developments in this field. The primary goal of the paper, as the authors put it, was to

provide an overview of CCA in relation to its foundation, approaches and applications. The first two of these characteristics are of particular relevance. In terms of its foundation, citation context in CCA is analysed in the full text of the paper, rather than in terms of simplistic citation counts and this analysis is both syntactic and semantic. The syntactic element is concerned with the location of the citation and its number of mentions (i.e. frequency). Here too, a distinction is made between formal (author and date) and informal (author only) citations. The authors suggest that location and frequency are influential factors which contribute to the significance of a citation. The semantic however, is concerned with how authors cite and why they cite and this is where the meaning and context of each citation are explored.

Omitted from Ding et al. (2014), however, as is the case with many other studies of CCA in LIS research is an illustration of how their approaches could be applied to authentic data. The approaches remain purely theoretical and on many occasions require a detailed level of discipline knowledge. The authors attempt to link their observations to linguistic analysis but do not demonstrate how this is achieved. Citation value and function seem to be studied as a means to automate analysis rather than for the sake of understanding the value of citations in the citing papers. This is probably due to the fact that LIS and linguistic studies differ in their goals. Finally, it appears to be the case that in its review of citation studies, this paper stresses context over content.

The next two studies are concerned with CCA and are purely quantitative. They are closely related to each other, with the first providing the foundations for the second, although they differ in focus. The first study, by Ding et al. (2013), applies text-mining algorithms to a dataset of 866 information science articles with the aim of shedding light on the contributions made by references. To do this, the authors considered the location at which the citations appeared across different sections of a journal article, and identified and compared highly-cited works using two different counting methods, ‘CountOne’ and ‘CountX’, representing each reference as mentioned once or x times, respectively. The major finding was that the citation rankings produced by CountOne and CountX differ. In other words, the number of times a bibliographic reference is cited in the citing paper is suggested to show the differential contribution made by that cited reference, which can subsequently be compared to other bibliographic references in the same paper based on their mentioning frequencies. Of course, this distinction overlooks the nature and contribution of each citation (cf. Amsterdamska and

Leydesdorff 1989). Simply put, the study examines citation contribution by looking at citation distribution and comparing highly-cited works using the above mentioned counting methods. The study uses the number of citation mentions as a means to identify and compare highly-cited works.

While the number of citation mentions is used to identify highly-cited articles in the above study, it is in turn employed to "quantify a scientist's impact" in Wan and Liu (2014: 2512). The authors conducted a quantitative analysis on a large-scale dataset of 10,305 papers. They propose a new version of the h-index (namely: 'WL-index') as an improvement on and replacement for existing bibliometric indices, arguing that it is capable of capturing research impact. They suggest replacing citation counts by citation mention counts in all the variant indices to capture citation mentions and demonstrate differences between individual researcher contributions.

In the last five years, LIS research has paid increasing attention to the use of computer-assisted analysis. Information scientists now use more techniques for the automatic identification and coding of citation functions. The main methods of retrieving citation functions are text-mining (Ding et al. 2013) and NLP and n-grams (as in Bertin et al. 2016). All the above studies conclude that devising a classification model and applying it using computer-assisted techniques is a non-trivial task. There is a consensus that human intervention is needed at one stage or another in the interpretation and validation of the quantitative results.

The recent trends reflect current thinking in LIS research on citation practices. They also demonstrate that the main challenges faced within this field are in the improvement of citation indexing and retrieval performance. A quantitative analysis has always been present in bibliometric research. While it is useful for offering an indication of the influence and reach of an individual's research, a quantitative analysis falls short of appropriately explaining the content and context of citations. This is where the role of textual analyses of citations becomes crucial in deciding on the function of citation.

Recently, LIS studies have begun to consider some useful aspects of citation analysis that have been long studied in linguistic research such as the relationship between the citing and the cited. Linguistic citation literature in turn has benefited substantially from developments in, and limitations of, LIS literature. The assumption that citations are unequal, following criticism of citation counting measures, was the starting point for

many linguistic studies of citing behaviour (e.g. Harwood and Petrić 2012), motivation (Brooks 1986) and functions (Petrić 2007; Harwood 2009). The next section (2.3) surveys the literature on citation practices from a purely linguistic perspective.

2.3 The citation process: linguistic research into citation

In contrast to LIS studies of citation, linguistic research emphasises the qualitative aspect of citation analysis. In general, linguistic research into citation practices considers citation as an important academic practice not only for the purposes outlined in LIS, such as assessment of research influence, but also for knowledge construction in academic discourse. Citation, in this view, is conceived as a linguistic phenomenon with syntactic and semantic variables (Jalilifar 2012). Linguistic studies of citation explore the nature and purpose of citing. Citation, as an academic practice, has received considerable attention in linguistic research in general, and in EAP research in particular (e.g. Swales 1986, 1990; Hyland 1999; Thompson 2001, 2005a, b; Charles 2006, 2007; Petrić 2007; Harwood 2009).

Over the last three decades, following criticism of LIS approaches, increasing attention has been given to the analysis of citations within and across various disciplines and genres. For example, there are studies of citation in research articles (Samraj and Monk 2008; Okamura 2008; Pickard 1995; Shooshtari and Jalilifar 2010), MA theses (e.g. Jalilifar 2012; Samraj 2008, 2013; Maher 2015), and PhD theses (e.g. Thompson 2001, 2005a; Ali 2016). The focus has been mostly on the analysis of citations in single disciplines (e.g. Pickard 1995 who analysed 11 articles on applied linguistics from the Linguistics and Language Teaching Journal, cited in Jalilifar 2012: 25), across disciplines or cultures (e.g. Thompson 2001, 2005a, b; Hyland 1999) or between sub-disciplines (Shooshtari and Jalilifar 2010). The above studies reflect variation in scope and methodology but in general, they are all concerned with citations from a purely linguistic perspective. That is, citations are analysed to understand how and why writers make certain choices and what influence those specific choices have subsequently on the development of an argument, or readers' perceptions of the cited materials (e.g. Martin and White 2005).

Before reviewing the linguistic literature on citation, it is important to define citation as a linguistic phenomenon. The process of citation has been studied under different labels and various terms have been adopted in the literature to refer to the broader

practice of citation (as used by Bazerman 1988; Swales 1990; Thompson and Tribble 2001; Thompson 2005a, b; Moed 2005; Hu and Wang 2014). One label, which is often used interchangeably with citation, is 'referencing' (e.g. Leydesdorff 1998; Coffin 2009). In his extensive review of citation theories, Leydesdorff acknowledges two views of citation. The first view belongs to LIS research into citation, whereby citation analysis is conducted as a gateway into the understanding of something else (e.g. the growth of science). In this sense, citation analysis is nothing but a tool for explaining or reflecting upon the accuracy of scientific growth measures. Leydesdorff (1998: 7) refers to citation analysis with such a purpose as "explanans", when citation is used to study something else.

The second view, which relates to linguistic research, is called "explanandum" (ibid: 7-8). In this view, citation is the linguistic phenomenon to be explained. Within this linguistic approach, Leydesdorff defines citations as "references to another [external] textual element, from the perspective of the citing article" (1998: 8, emphasis added). Not only does Leydesdorff effectively define citations, he also includes the writer's perspective in the identification of citations. In linguistic literature, citation is a fixed concept which always includes a reference to some external source(s) in academic prose.

Within the study of citation, there are other fixed grammatical items (e.g. 'projection' as within the SFL tradition, see Halliday 1994, 2004; Forey 2002: 172-186 on the use of 'projecting' [reporting] and 'projected' [reported] clauses) or grammatical concepts (e.g. direct versus indirect speech and reporting as in Thompson and Ye 1991; Thompson 1994). The latter are very traditional grammatical terms and may be found in Quirk et al. (1985). Different, yet relevant, to citation are more abstract concepts such as discourse representation and/or intertextuality (Fairclough 1992) as well as heteroglossia (Bakhtin 1934-1981, cited in Coffin 2009: 169). Other concepts are much less fixed such as attribution (e.g. Sinclair 1987; Hunston 1993, 1995, 2000; Hyland 1999, 2000; Petrić 2007; Dontcheva-Navratilova 2016; see Section 2.3.1 below for a discussion of attribution). These terms reflect terminological and, to varying degrees, ideological differences. Each of the aforementioned terms has its own formal characteristics and differs from citation in some aspects. For example, indirect speech in traditional grammar coincides in some respects with citation, and differs in others. The former is used more in news reports or novels whereas citation is a common feature

of academic writing. In sum, there are different ways of identifying the target in each concept. In this study, however, the focus is only on two closely related terms: citation and attribution.

Central to any analysis of citation are the notions of 'averral and attribution' (Sinclair 1987) which are discussed in the following sub-section. Under the current section (2.3), I review the most relevant themes in linguistic research into citation. One important theme in citation analysis is related to formal typologies that are based on syntactic division or the correlation of tense, voice and aspect. This strand of research is reviewed in Sub-section 2.3.2 (Citation and form). Sub-section 2.3.3, then, investigates citation literature in terms of stance and evaluation. The last sub-section (2.3.4) reviews literature on citation functions.

2.3.1 Averral and attribution

In the 1980s, John Sinclair (1985, 1986a, 1987, 1988) pioneered research into whose textual voice is being expressed in the text. Sinclair identifies two notions that are important in the construction of knowledge: these are *averral* and *attribution*. In a seminal paper on the description of written texts, Sinclair (1987) treats writing as an interactive skill. He pays specific attention to the writer-reader interaction and notes that "the competent writer is sensitive to his readers" (ibid: 1) and that the value of a text may be measured by the extent to which it has an impact on the readers.

To account for this interaction, Sinclair introduces the notions of averral and attribution as two forms in knowledge construction. In its simplest form, averral is defined as "the default condition" of a written, or spoken, text (Tadros 1993: 101) in the sense that an averred statement identifies "the textual voice of the writer" (Thompson 2012: 121). In averral, the writer accepts full responsibility for the proposed information. In attribution, on the other hand, the writer uses intertextual markers signalling "antecedent authorial voice" (e.g. in the form of citation in written academic discourse, see Thompson 2012: 121) whereby part, or the whole, of the statement in which the attribution occurs is positioned to belong to an external source. The writer thus transfers responsibility for an attributed proposition to its originator. Both averral and attribution are crucial in the identification of textual and intertextual voices (Groom 2000).

Averral is used in one way in the literature, which is to refer to text or part of the text that belongs to the writer (or speaker). However, the term 'attribution' is problematic

as it is used in different ways. As a result of the different uses of this term, it has different relations to the notion of citation. In a broad sense, attribution has been studied either as a form of interaction in discourse (as in the Birmingham school led by Sinclair 1987 and followed by Hunston 1995, 2000; Groom 2000; Thompson 2012) or as a function of citation (e.g. Petrić 2007 and later in Gol et al. 2014; Hu and Wang 2014; Dontcheva-Navratilova 2016). However, the view of attribution as a function of citation is irrelevant to the discussion of averral and attribution here. This is because the functional view of attribution (e.g. Petrić 2007) does not include instances of averral. In that view, it is taken for granted that there is a citation and that attribution is included as one function of citing. Therefore, the use of attribution by Petrić (and those who adopt her typology) is accidental and relates to functional analysis of citation (see Sub-section 2.3.4 in this chapter for a discussion of functional models including that of Petrić 2007). Furthermore, Petrić's use of attribution is similar to some extent to the sub-system 'Attribute' in Martin and White (2005), which is reviewed in Sub-section 2.3.3 in this chapter.

Sinclair's perception of attribution as a form of interaction has been highly influential in the literature on writer-reader interaction. For example, Hyland (1999, 2000) conceives attribution and reporting as two important realisations of "the academic writer's concern for interactions with an audience" (Hyland 2000: 20). Hyland (2000) analyses citation practices in eight academic disciplines and introduces citation as one of the tools available for writers when producing new knowledge claims. When accounting for attribution and its relevance to citation, Hyland is particularly effective in explaining the social function of explicit references to external texts. He views attribution as a social phenomenon which enables the writer to "establish a persuasive epistemological and social framework for the acceptance of their arguments" (2000: 22).

In order to account for the different meanings generated by attribution and their relevance to citation, it is important to distinguish between citation and attribution as two overlapping concepts. To start with, as will be shown below, attribution as a linguistic phenomenon co-exists with averral. Every statement contains elements of averral, attribution or both (although averral and attribution may be unsignalled, see Examples 2.4 and 2.9 in this sub-section). Citations, whether integral or non-integral, may contain attributed information to other external sources or may simply report

previous research. To demonstrate the overlap between citation and attribution, consider the following examples obtained from the data of the current study (the list of papers analysed in this thesis is included at the end of the reference list):

Example 2.1:

> Adger and Barnett (2005) suggest that encouraging migration as a solution to climate change detracts from … [2009_Mortreux and Barnett]

Example 2.2:

> Jodha et al. (2000) used 50 indicators of … [1995_Millette et al.]

Example 2.3:

> Utilitarians argue that … [1995_Beckerman and Pasek]

The first two examples are citations of Adger and Barnett (2005) and Jodha et al. (2000), respectively. Example 2.1 exemplifies indirect speech that is attributed to Adger and Barnett. Example 2.2, however, comprises a report of Jodha et al.'s research and includes no attribution to the cited authors (see the use of 'action', as opposed to 'mental' or 'verbal', verbs of reporting in Bazerman 1984). Example 2.3 is another example of attribution but, unlike 2.1, contains no citation or specific reference; instead, the writers in 2.3 attribute the information to a school of thought (Utilitarians). Generally, in published academic writing, attribution to unidentified sources is deemed bad practice.

The literature on attribution focuses on examples such as 2.1 and 2.3 above, and excludes examples such as 2.2 (e.g. Sinclair 1987; Hunston 1995, 2000). In that view, attribution can be cited or non-cited (e.g. 2.1 and 2.3, respectively). At the same time, citation can also be seen as the broader umbrella term that includes reports of speech or thought (attribution, as in 2.1) and reports of research (2.2). The overlap between citation and attribution is diagrammatically represented in Figure 2.1 (overleaf).

The shaded part (highlighted in blue) in Figure 2.1 is the area of overlap between citation and attribution. This area includes instances which can be coded as both attribution and citation (e.g. 2.1). However, as the diagram shows, there are instances which belong either to attribution (e.g. 2.3) or citation (e.g. 2.2). In the present study, I

analyse all instances of citations, whether or not they exhibit attribution (see Chapter 3).

Before I investigate the interplay between text averral and attribution, I shall first exemplify these terms. Text averral is usually "unmarked" in the text (Tadros 1993: 98, cited in Jabbour 1997: 140) and is, by default, attributed to the text producer. However, there are occasions of 'emphasised averral' (Hunston 2000) where the writer emphasises his/her assertions. By way of comparison, consider the following examples (also taken from the current thesis corpus):

Example 2.4:

> The Queensland landscape has undergone steady transformation since European settlement, with deforestation accelerating since the 1960s. [2009_McAlpine et al.]

Example 2.5:

> We argue that such industry and government agendas are not environmentally responsible … [2009_McAlpine et al.]

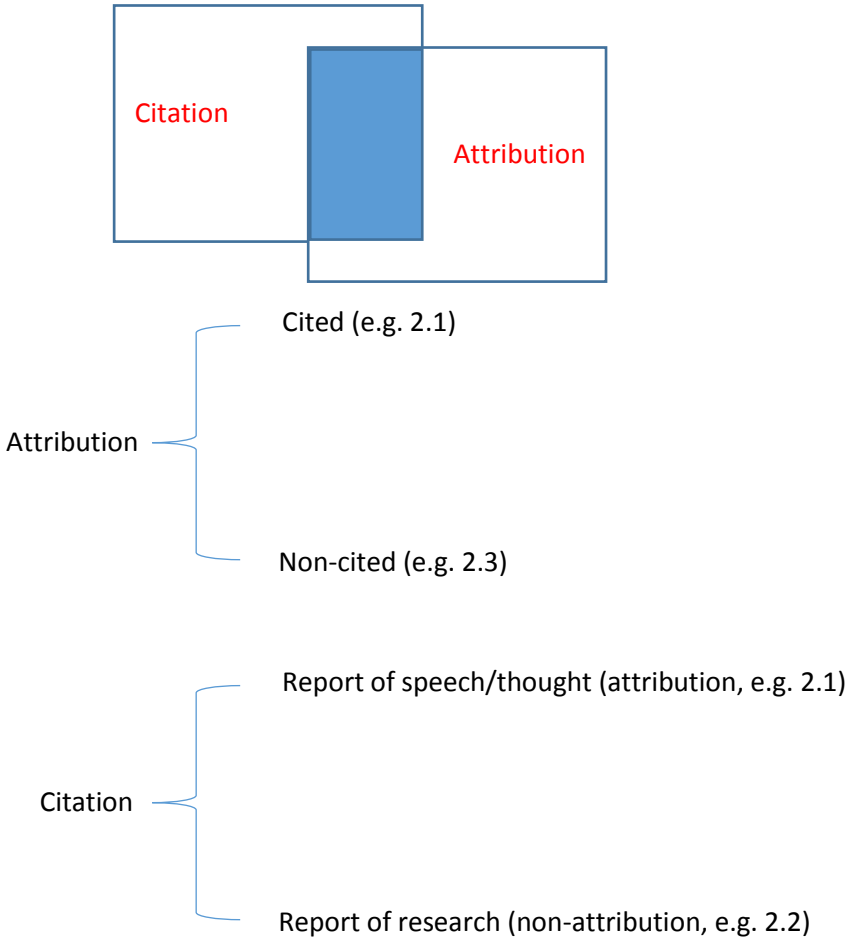

Figure 2.1: Citation and attribution as two overlapping concepts

Both examples are averred by the same authors (McAlpine et al. 2009) and include no citation. However, in 2.4, the authors do not signal their involvement, unlike in 2.5. Put differently, the authors present the information in Example 2.4 as a given fact (*has undergone steady transformation*). The choice of tense and aspect plays an important role in the construction, presentation and interpretation of the proposed information as a given fact (Hunston 1993, 1995, 2000). Thompson (2012: 121) notes that via such choices the writers accord the statement "the status of 'fact'". The present perfect use in Example 2.4 contributes to the presentation of the information as a fact (i.e. as something which has happened).

In Example 2.5, the same writers highlight their contribution and emphasise their role in the argument. This statement exemplifies "averred assessment" (Hunston 2000: 187). The authors retrospectively evaluate the *industry and government agendas* which appear in previous text. The anaphoric reference is achieved by the use of *such*. The writers make their involvement in the negative evaluation of those agendas very prominent via the choice of *We argue*. The authors could have passed their negative evaluation of the agendas without the insertion of the clause *We argue that* (cf. Hunston 1995, 2000, 2004) so that the same statement starts with *Such industry and government agendas …*. However, they chose to emphasise their role in the assessment; therefore, this is a clear example of emphasised averral. Signalling the writer's involvement in thematic position enables the writer to give prominence to their viewpoint and evaluation of the information in the remainder of the statement, in instances of averral or attribution. Forey (2002: 173) writes:

> It is suggested that the writer, through the use of projecting clause in thematic position, is able to use viewpoint as the starting point for what will follow. Projection will affect the way in which the reader interprets the following and perhaps even the preceding clause and to some extent beyond the clause.

As noted by Perez-Llantada (2009), writers can emphasise their personal commitment in averred statements using one of at least three common constructions: 1) 'I'-as-subject pattern, 2) 'we'-as-subject pattern and/or 3) agentless constructions such as passives or 'it'-as-subject patterns. In the first and second constructions, the distance between the writer and the averred proposition is very close, while in the third construction the

distance increases making the role of the writer less prominent. For example, the writers signal their involvement more prominently in instances such as 'I argue' or 'We strongly believe' more than in other constructions such as 'It is well-known that'. However, it could equally be argued that some 'it' patterns (e.g. It is well-known that) indicate shared knowledge and thus minimise the social distance between the interlocutors.

Groom (2005) and Molino (2010) analysed ''it'' patterns in active and passive forms of attribution. They conclude that the impersonal 'it' allows writers to comment on the attributed information in various ways, for example on the importance, adequacy and difficulty of the attributed propositions. The writer's choice reflects the degree of alignment with the cited information, as Hyland (2009) argues.

As noted above, attribution is linked to "the use of a manifest intertextual marker to acknowledge the presence of an antecedent authorial voice" (Groom 2000: 15). This explicit marker is usually a reporting verb or noun (Hunston 2000) that functions to inform readers that the responsibility for the attributed proposition is transferred "from the current writer to [an] attributed author" (Sinclair 1981: 78). Via the use of certain reporting verbs or nouns (e.g. *demonstrate, disclosure*, see Hunston 2000), the writer can also reclaim responsibility for an attributed proposition. Hunston (2000) asserts that the choice of reporting verbs and nouns enables academic writers to manipulate averrals and attributions, so that the responsibility for an attributed piece of information is either 'relegated' or 'reclaimed' by the writer. Likewise, Groom (2000) notes that there are varying degrees of shared responsibility, depending on how the writer presents the attributed information.

Substantial body of the literature has been directed towards the correlation between averral and attribution on the one side and writers' stance and responsibility on the other (e.g. Hyland 2000; Charles 2003; Martin and White 2005; Thompson 2012; Adel and Garretson 2006; Holmes and Nesi 2009). The majority of this research has focused on the use of reporting verbs and their role in unpacking the writer's stance towards, or evaluation of, the attributed information (e.g. Thompson 1994, 1996; Hunston 1993, 1995, 2000; Thompson and Ye 1991; Hyland 2000). Consider the following examples:

Example 2.6:

Gusterson (1996) argues that nuclear weapons scientists are taught to value science and technology and to trust humans' ability to control it in socially beneficial ways. [2008_Lahsen]

Example 2.7:

O'Collins points out that emigration places severe strains on immigrants, host communities, and on those who remain behind.[fn] [1992_Pernetta)

Example 2.8:

… *As Bauman writes*, "Thatcher's fanciful imagination (became) a fairly precise description of the real world, as seen from the *inside* of its inhabitants' experience." (REF) [2009_Rathzel and Uzzell]

Example 2.9:

Changes in species distributions have also been observed in both above-ground (REFs), and below-ground biotic communities (REFs). [2010_Hagerman et al.]

The verb *argues* is a clear signal of attribution in Example 2.6 as it neutrally indicates to the reader that what follows belongs to the cited author(s). No evaluation is given by the citing writer of the reported proposition, which is underlined. In Example 2.7 and 2.8, the writers also attribute the underlined propositions to the sources but, unlike 2.6, they reclaim responsibility for the attributed information. The choice of the verb *points out* and the structure *As … writes* makes it clear that both citations are *endorsed* at the proposition-level. Hunston (1995: 135) makes it clear that the choice of the reporting verb *point out* construes the writer "as committed to the truth of the attributed statement". Likewise, the choice of the structure *As Bauman writes* in 2.8 signals the writers' *endorsement* of Bauman's argument. Although the underlined quote in 2.8 is attributed to Bauman, the writers share responsibility for the quoted proposition. There is evidence in the literature that writers can retain "a strong authorial presence" via the integration of direct quotations into their own averred statements (Thompson 2012: 124). 2.8 is an example of shared responsibility whereby the source's argument is integrated into that of the writers'.

In Example 2.9, the changes in species distributions are originally attributed to, and observed by the sources. They are presented as reported facts. Hunston (2000) effectively makes the distinction between reported 'facts' and 'assessments'. She notes that the primary indication lies in the implication they portray to the reader. "Facts will

not be further evaluated for their truth-value", unlike assessments (ibid: 188). In this sense, the writers not only acknowledge the sources which documented those changes but also wish to provide background evidence for the proposed information, for which they also share responsibility.

We now turn to the interplay between the notions of averral and attribution as discussed in the literature (e.g. Sinclair 1987; Hunston 2000; Thompson 2012). Sinclair (1985, 1986a, 1987, 1988) maintains that "every attribution is embedded within an averral" (cited in Hunston 2000: 179). To demonstrate this, consider Example 2.1, represented here as Example 2.10:

Example 2.10:

> Adger and Barnett (2005) suggest that encouraging migration as a solution to climate change detracts from the need for adaptation policies to allow people to "lead the kind of lives they value in the places where they belong" (p. 328) [2009_Mortreux and Barnett]

There are two propositions in the above example:

> 1) encouraging migration as a solution to climate change detracts from the need for adaptation policies to allow people to "lead the kind of lives they value in the places where they belong" (attributed to the source)
>
> 2) Adger and Barnett (2005) suggest that (averred by the writers)

The first proposition is modified by the reporting verb *suggest*. This verb indicates that the writers are uncommitted to the attributed proposition (Hunston 1995), thus detaching themselves from its truth value (Tadros 1993). The following is another representation of the same example, as seen by Sinclair: The writers (Mortreux and Barnett 2009) aver that the source (Adger and Barnett 2005) aver (suggest) that encouraging migration as a solution to climate change detracts from the need for adaptation policies to allow people to "lead the kind of lives they value in the places where they belong".

The division proposed by Sinclair has informed many studies to demonstrate the interplay between averral and attribution (e.g. Tadros 1993; Hunston 1995, 2000; Jabbour 1997; Groom 2000, 2007; Hyland 2000; Hyland 2009; Thompson 2012;

Jalilifar 2012; Maher 2015, 2016). A considerable body of this and other research has addressed students' writing (in undergraduate student writing as in Lee 2010; in MA theses as in Samraj 2008; Basturkmen 2009; and PhD theses as in Thompson 2001, 2005a, b, 2009, 2012) and related aspects such as plagiarism (e.g. Luke and Kearins 2012). Other genres have been investigated as well (e.g. in news discourse, as in Jullian 2011; or in research articles as in Berkenkotter and Huckin 1995; Jabbour 1997; Hyland 2000).

In sum, the study of attribution has stressed sentences or parts of sentences which indicate the writer's stance towards the cited material (e.g. Sinclair1985, 1987; Tadros 1993; Hunston 1995; Coulthard 1994). Later research has stressed the writers' voice in averral and attribution (e.g. Groom 2000, 2007; Thompson 2012). As for the combination of attribution and stance, extensive research has been conducted on the role of reporting verbs in the expression of the writer's stance (e.g. Thompson and Ye 1991; Thompson 1994, 1996; Hyland 2000; Hunston 1995, 2000, 2004; Thompson 2012). Other research has stressed the use of specific lexical items (such as the node 'that' in Maher 2015) in postgraduate law students' writing in three universities. The major goal was to investigate how students use the node 'that' in university assignments and MA theses. Maher combines many aspects in the analysis of different constructions and uses concordance lines with 'that' as the key word. By doing that, he was able to identify the frequency of various features such as tense, subject position and stance. The main finding from Maher's study is that it is difficult for PG students to successfully manage averral and attribution at the proposition-level.

The overall conclusion from the averral and attribution literature is that both forms are closely related and that their management is critical for successful knowledge construction (Hunston 2000). There is also a consensus that student writers should be exposed to the interplay between the two forms, to make them aware of their choices in comparison with the choices of published authors (e.g. Maher 2015). Most of the literature, however, has stressed the interplay at the proposition-level despite Sinclair's acknowledgement of their roles extending beyond the proposition level (refer to the analysis of Example 1.1 in Chapter 1).

In this thesis, I adopt Sinclair's view of attribution as a form of interaction. I look into instances of cited attribution (i.e. where a text writer uses "manifest intertextual

markers" in citations – see Groom 2000: 15; Thompson 2012: 121 - to report speech or thought). I also include reports of research as in non-attributed citation. In citations, the writer identifies one or more external source(s). Depending on the way in which the writer presents the cited research, they can relegate or reclaim responsibility for that research.

2.3.2 Citation and form

Since the 1980s, a substantial body of literature into citation practices has emphasised the correlation between tense, voice and aspect in reporting (Sub-section 2.3.2.1). Another strand of research has investigated different forms of reporting, following Swales' (1990) division of citations into either integral or non-integral (Sub-section 2.3.2.2). Identifying tense, voice or form in reporting is a straightforward, objective process in which the writer's subjectivity plays no role (unlike the identification of a writer's intentions or reasons behind choosing a specific tense or aspect).

2.3.2.1 Tense and voice in reporting

Research of this kind investigates the use of, and the meanings carried by, different tense, aspect and voice combinations (e.g. Oster 1981; Riddle 1986; Malcolm 1987; Gunawardena 1989; Pecorari 2013). In the identification of tense and voice choices, the analyst's subjectivity plays no role. However, interpreting the reasons behind these choices is a subjective matter. Only the text writer can identify with certainty the reason or motivation for using certain tense or aspect when reporting (though there are recall issues, as reported in Harwood 2009). In this view, generalisations about the use and reasons behind tense choices could be seen as, at best, speculations over the writer's intentions behind such choices.

The argument that tense choice in scientific writing is important first appeared in Lackstrom et al. (1972, cited in Thompson 2001: 101). Since then, tense choices have been linked to stance (e.g. Swales 1990) or evaluation (Thompson and Ye 1991; Hunston 1995, 2000), as well as function of reporting (e.g. Pecorari 2013). By function, the literature refers to the reasons for using a particular tense and voice to refer to previous research (e.g. the use of present perfect in non-integral passive citations to introduce sub-topics or state generalisations to be elaborated on, see Swales 1990, cited in Thompson 2001: 102). The past and present are the main tenses examined since the focus was on the use of reporting verbs and structures. Table 2.1 (overleaf) summarises

the most frequently reported reasons for using the following three tense and aspect combinations: past simple, present perfect and present simple. It also gives an indication of the type of research that is reported in each tense and/or aspect (e.g. evidence, assessment or interpretation).

As important as the indication of tense choices in the literature is the indication of the shift between tenses (Sinclair 1985; Jabbour 1997). In general, the present tense is used to introduce external views and/or to comment on them (e.g. in the use of mental reporting verbs such as *believe* or *claim*) while the past tense is used to report previous research (research verbs such as *conducted* or *analysed*). Sinclair (1985: 23) effectively argues that the change in tense, or even aspect, "allows a change in posture". He argues that it is possible to identify organisational divisions within a text marked by what he terms posture change. Sinclair highlights different ways in which posture changes: the shift between tenses and/or aspects when reporting is one way and attribution is another.

Jabbour (1997) examines attributions in medical research articles following Sinclair's argument. Jabbour views tense as an element of a writer's evaluation and text organisation. She notes that tense is "an organising device between the interactive and the autonomous segments, and consequently is a marker of posture change" (1997: 146). She argues that the structure 'has/have been shown to be' performs the function of relating the past to the present. Prior to that, Oster (1981) observes that such combinations of tense and aspect reflect the writer's decision to relate, and possibly align, previous research to their present research. It is important to note that, unlike tense choice identification, the analysis of the implication of the shift in tense and aspect requires some level of interpretation. The analyst is involved in constant attempts to deduce the reasons for that shift and its consequence on the build-up of the argument, in addition to the writer's position and text organisation.

In reporting, the use of the past simple, present perfect and present simple has been compared in different genres and disciplines. Table 2.1 summarises the general findings from some studies with respect to these three tense and aspect combinations. However, detailed discussion of the indication of each tense choice falls beyond the scope of the current thesis.

Overall, three observations can be made from the literature on tense and aspect combinations (e.g. from Shaw 1992; Thompson 2001 and recently in Pecorari 2013).

First, the choice of tense in integral or non-integral structures can be linked to the writer's stance and degree of generality or alignment. Second, these studies, considering the small corpora they have examined, reflect tendencies and so are not

Table 2.1: Indication of tense (and aspect) choices in reporting as they appear in some studies

	Past simple	Present perfect	Present simple
Lackstrom et al. 1972	**- reporting unimportant/old-fashioned research - citing individual studies of sub-topics - lack of generality**	**- reporting past research with current relevance - introducing sub-topics - generalisation about reported events**	**- generalisation of claims**
Oster 1981	**- referring to quantitative results of non-supportive nature - claiming non-generality of past literature**	**- indicating continued discussion - claiming generality of past literature**	**- referring to quantitative results of supportive or irrelevant research, rather than discussion of previous literature**
Riddle 1986 (on the use of past tense in ESL context)	**- "true in the past … the notion of completion is contextually based implication" (p. 268) - past association of a situation – "indicator of subjective attitude" (p. 269) - to organise discourse and background information**		**- describing past events (e.g. Smith (1990) writes …).**
Malcolm 1987 (also cited by Thompson 2001)	**- referring to specific experiments**	**- referring to areas of enquiry**	**- claiming generalisation**
Gunawardena 1989 (on the use of the present perfect in Biology and Biochemistry Research articles)	**- when focus is on event - reporting irrelevant research - to cite individual studies**	**- to introduce a number of studies on a particular topic - reporting past research with current relevance or to signal continuity of research which began in the past - when focus is on consequences or current state following a past event - repeated action in the past - an action which took place in the**	

		"immediate past" (p. 267) or that which is expected to be continued in the future	
Swales 1990	**- implying social distance between citer and citee**	**- implying closeness to the study (cited in Thompson 2001: 102)**	**- increasing proximity between current and previous texts**
Shaw 1992	**- not for generalisations** **- reporting irrelevant research**	**- reporting relevant research**	**- reporting as factual**
Swales and Feak 2004	**- reporting past events and single studies**	**- reporting areas of enquiry**	**- state-of-the-art knowledge** **- relevant research in student writing**
Pecorari 2013	**- reporting past literature**	**- reporting past literature that is still valid**	**- to show previous work as still valid and endorsed (e.g. X shows)**
Maher 2016 (in post-graduate legal writing)	**- to report court decisions**		**- "to report the contents of statutes" (p. 106) or academics' views** **- to cite an exact location in "a legislative document" (p. 96)**

general rules in reporting. Finally, the change of tense and aspect is important in the analysis of writer evaluation, text organisation and citation function. By function, here, I do not mean the reason or intention behind citation, as in the literature; instead, I mean its rhetorical function as observed in the text (i.e. the role the citation plays in the argument that the writer is building). There is a tendency in the citation literature to treat the writer's intentions behind citing as synonymous with the citation functions (see Section 2.3.4 in this chapter for a discussion of the literature on the rhetorical functions of citing). Most of these studies conflate the writers' perspectives (their intentions, motivations or reasons for citing) with the analyst's or readers' perspectives on the citation function. The current thesis makes a clear distinction between citation functions on the one hand, and the writer's intentions or motivations on the other hand. The function of a citation relates to how that citation contributes to the writer's argument as identified by textual cues in the block. The writer's intentions or motivations are their reasons for citing and can only be accounted for by the citing writer.

2.3.2.2 Swales' (1990) formal typology

Swales (1981, 1986, 1990) furthered the investigation of tense and aspect in citational statements. He argues that tense selection in citation can be indicative of the writer's "stance" towards the cited authors or their findings (Swales 1990: 154; also cited in Thompson 2001: 102). The choice of tense and aspect, combined with strategic choice of reporting verbs and/or structures enables writers to convey their stance and commitment towards cited propositions (Swales 1990: 151). Following the same path, Hunston (1993, 2000, 2004) notes that the choice of tense/aspect and reporting verbs (e.g. *has shown*, see Hunston 2000: 179) enables writers to reclaim responsibility for a reported argument. Some reporting verbs such as *show*, *point out* and *prove* (see Thompson 1994, 1996 for a comprehensive list) clearly signal the writer's *endorsement* of, and commitment towards the reported information unlike other verbs such as *propose* and *suggest*, which indicate that the writer is undecided about the truth of cited propositions and therefore withdraws from sharing responsibility. In terms of tense, Swales argues that the shift from the past to the present tense indicates increasing proximity between current and previous texts, and also between the writer and reader. Consider the following invented examples:

Example 2.11:

Smith (2000) pointed out that the destruction of Rome has caused the war.

Example 2.12:

Smith (2000) points out that the destruction of Rome has caused the war.

There is one difference between the above examples: that is, the tense. Following Een (1982) and Ard (1985), Swales suggests that Examples such as 2.11 place the cited work in a 'historical context' (p. 154) while Examples like 2.12 emphasise relevance to the (current) citing text. The correlation between tense and reporting verbs has long been discussed in the literature (see, for example, Thompson and Ye 1991; Thompson 1994, 1996; Hunston 1989, 1993, 2000).

Within the field of applied linguistics, Swales examines tense choices and stance indication as part of a broader model which investigates rhetorical movement in article introductions. His first attempt to account for this movement was in 1981 when he introduced a '4-move' model for article introductions. However, as he concedes (1990:

140-148), that attempt did not stand the test of critical commentary (mainly by Jacoby 1987). Several studies have criticised his moves and the difficulty of drawing the boundaries between the moves (e.g. Lopez 1982; Crookes 1986). Therefore, Swales offered a modified version of his model in 1990 which he claims "captures a number of characteristics of RA [research article] introductions" (140). He refers to the revised version as the 'Create a Research Space (CARS) model'. Swales argues that researchers use references in order to "open a rhetorical space" (Charles 2004: 157) which they can then occupy. This role of citations corresponds to some roles identified by Dudley-Evans (1986: 135) in MSc dissertation introductions. Swales' (1990) investigation of RA introductions resulted in the identification of three moves and a number of steps within each move (see Figure 10, in Swales 1990: 141, shown here as Figure 2.2). Once the revised model was formulated, Swales tested his model and found a preference within some articles to start with certain moves. For example, some of the RA introductions he examined started with Move 3 (see Swales and Najjar 1987; Swales 1990).

In his CARS model for article introductions arising from an ESP tradition, Swales (1990) notes that at some stage where writers establish the territory for their research, they need to, and invariably do review items of previous research (see Move 1, Step 3 in Figure 2.2). During this process of reviewing the literature, and writing up the results of their reviewing process, writers employ citations. At this stage, writers need to link *what* has been found or claimed with *who* found/claimed it (ibid: 148). To achieve this goal, writers need to provide a *specification* of previous research, an *attribution* and a *stance* towards the sources or their claims (ibid: 148).

Swales' analysis of RA introductions resulted in the identification of two basic forms of citations. These are 'integral' and 'non-integral' forms (1990: 148). In integral citations, the name of the cited source occurs as "some sentence-element" of the citing statement (p. 148). Swales (1990) argues that the source name can occur as a subject of an integral citation as in this present statement (*Swales (1990) argues*). It can also occur as part of a possessive noun phrase (e.g. Smith's (2000) finding …), as a passive agent (was found by Smith (2000)) or as "an adjunct of reporting" (to use the term given by Tadros 1985, also cited in Swales 1990: 148) as in the next statement. According to Swales (1990), tense and reporting verb choice in integral citations can be indicative of the writer's stance towards the cited material.

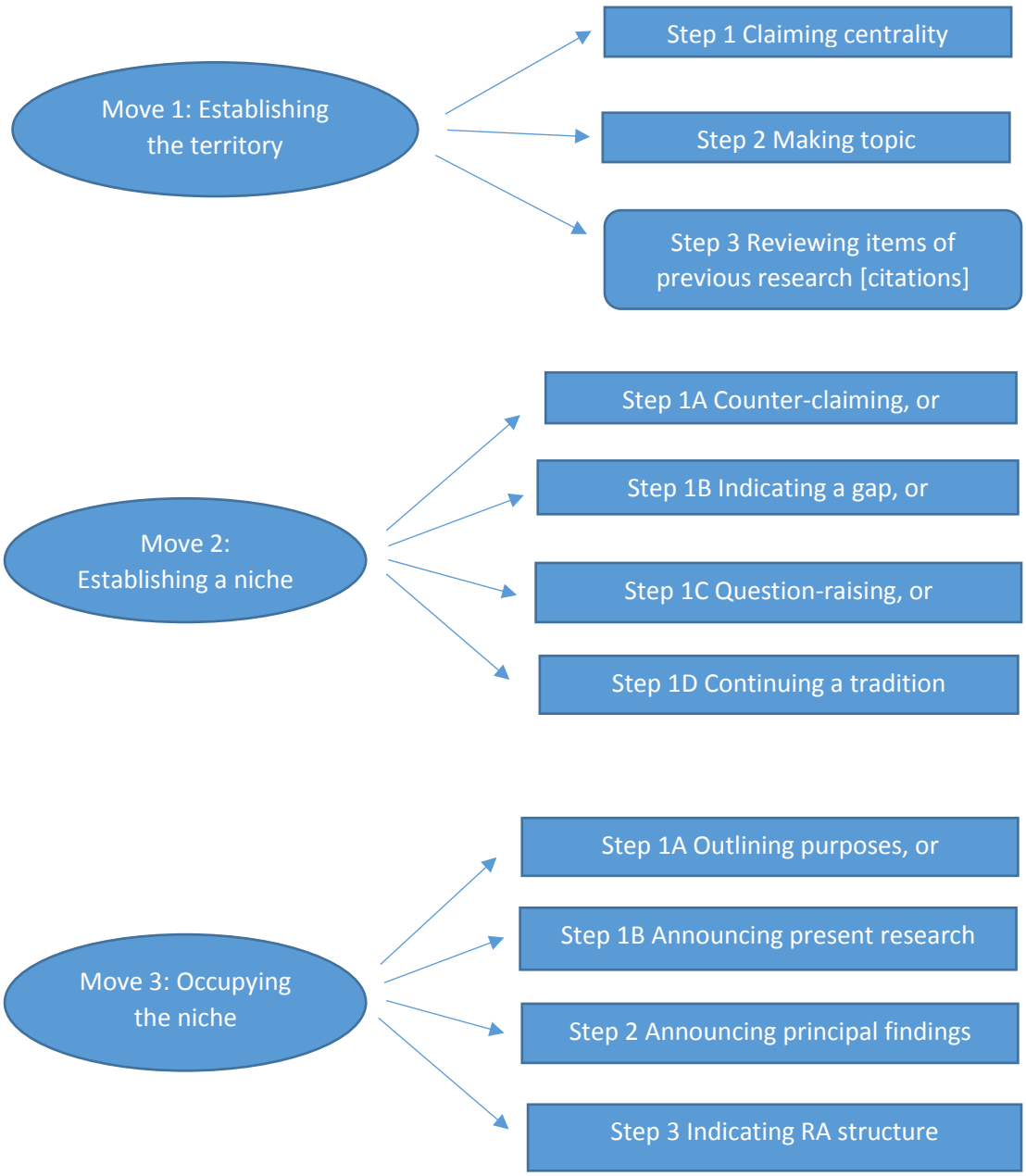

Figure 2.2: The revised CARS model for article introduction (Swales 1990: 141)

The second form of citation is the non-integral form. As the name suggests, in contrast to integral forms, the name of the researcher plays no grammatical role in the construction of the citing statement. It appears instead as external element, usually indicated in parentheses or referred to by a superscript number. While the appearance of the writer's name in integral forms (e.g. as subject in Theme position) signals which parts are averred and which are attributed, the case can be different in non-integral forms. In other words, positioning the name of the source at the end of a statement (see Example 2.18 below) occasionally makes it unclear which parts of the statement are

averred and which are attributed (see the analysis of Example Nd in Table 2.2 below). Therefore, identifying the boundary between text averral and attribution (and also responsibility for certain lexical items) can be problematic. In this view, the analysis of non-integral forms can be a complex task. This could be the reason why it has not received much attention in the literature compared to integral forms.

According to Swales (1990), non-integral forms are usually associated with non-reporting citations. The distinction between reporting and non-reporting was first raised by Bazerman (1984) and reemphasised in Swales (1990). Bazerman (1984: 177) traced the textual development in RAs from the *Physical Review* in the period 1893-1980. He distinguishes 'mental' and 'active' verbs that are used in references. He found an increase in the use of *active* verbs (e.g. *used*) and a decrease in mental verbs (e.g. *believe*). This verb choice resulted in a relative trend from reporting to non-reporting citations. Swales convincingly attributes this trend to the Physical Review's use of numerical/superscript system, which usually does not tolerate integral reporting structures (e.g. Ref. 2 argues).

Swales limits his study to citations which include the name of the researcher with the date of publication. Cases of general reference to schools of thought or even to researchers with no specifically identifiable source are treated as topic generalisations rather than as citations, even though they might be important in other aspects of the study of attribution. This is because it is sufficient under the terms of his model to consider only proper cases of citations (i.e. with reference and date, either specified or designated as being in the future, as in *in press*). Consider the following examples taken from Swales to demonstrate the difference between cases with and without proper citation (shown as examples 2.13 and 2.14):

Example 2.13:

> Chomsky and his co-workers (e.g. Napoli, 1988) have recently modified their position.

Example 2.14:

> Chomsky and his co-workers have recently modified their position.

In 2.13, the non-integral source appears as an example of Chomsky's co-workers. In 2.14, there is no specific reference to a source. Only the first version (e.g. 2.13) above

is treated as a citation under Swales' model and in this thesis as well. Swales' analysis of tense and aspect, however, is still applicable to examples such as 2.14.

Within the analysis of citation forms, Swales makes the distinction between reporting and non-reporting citations (see also Bazerman 1984 and Hunston 1995, *Report* versus *Attribution*). Under reporting, Swales includes all citations with a reporting verb such as *show* (see Examples Ia and Nb in Table 2.2 below). Examples without a reporting verb are classified in his model as non-reporting (e.g. Ie and Nd).

Table 2.2: Reporting and non-reporting citations in integral and non-integral forms, according to Swales (1990: 149)

	Integral	Non-integral
Reporting	Ia Brie (1988) showed that the moon is made of cheese.	Nb It has been shown that the moon is made of cheese (Brie, 1988).
Non-reporting	Ie According to Brie (1988), the moon is made of cheese.	Nd The moon is probably made of cheese (Brie, 1988).

One advantage of Swales' model is that it is "easily applicable" (Swales 1990: 148) as it relies on the identification of explicit linguistic features at the surface of the citing statement, such as the location and contribution of the source. Therefore, it is an objective process. The disadvantage, however, is that the model stresses citation syntactic form at the expense of reporting verbs or structures. Swales' view of non-integral citations with no reporting verbs as non-reporting does not seem to be accurate on some occasions. In Ie, for example, the structure *According to Brie* indicates that the proposition *the moon is made of cheese* is attributed to, and reported from, Brie. Brie is therefore responsible for the cited proposition. Also, in Nd, one possible reading is that the information was proposed by the source (Brie). *Probably* in Nd may be averred or attributed.

Swales (1990: 150) admits that his dichotomous classification does not work well in the face of uncertainties associated with some verbs (e.g. *find* and *be associated with*).

Yet, before one embarks on the analysis of (non-)reporting verbs, they should consider the source of the information. The source in Nd is there for a reason, regardless of the uncertainty associated with *probably*. Overall, it seems that Swales relates his view of reporting/non-reporting to the role of the cited source in the construction, more than the interpretation, of the citing statement. In other words, if the cited source is explicitly introduced (e.g. via the insertion of a reporting verb) as the originator of the cited argument, then the citation is reporting. On the other hand, if the cited source is seen as an intermediary source, with the possibility that it also cites the information from another source, then the current citation is non-reporting. However, this issue is highly subjective. Consider the following two interpretations of Nd:

1- Brie argues that the moon is probably made of cheese. (reporting)
2- The argument that the moon is probably made of cheese is cited in Brie but [had been] argued by someone else. (non-reporting)

Both interpretations of Nd are possible, although the second may be uncommon or unlikely. Swales classifies all non-integral citations with no reporting verbs (e.g. Nd) as non-reporting. However, I argue that the reader's interpretation, or knowledge of the source arguments can play a role in viewing such citations as either reporting or non-reporting. Swales treats modals (together with present tense choices, as in 'The moon *may be made of* cheese (Ref).') as signals of non-reporting. However, hedges and uncertainties play a broader role than merely signalling citations as reporting or non-reporting. As noted in the previous Sub-section 2.3.1, we need to distinguish reports of speech (reporting) from reports of research (non-reporting). Consider the following four invented examples:

Example 2.15:

> Smith (2000) argues that the impact of deforestation on the environment is unprecedented.

Example 2.16:

> Smith (2000) found that the impact of deforestation on the environment is unprecedented.

Example 2.17:

> Smith (2000) examined the unprecedented impact of deforestation on the environment.

Example 2.18:

The impact of deforestation on the environment is unprecedented (Smith 2000).

Example 2.15 is clearly reporting because of the presence of the reporting verb *argues*. There are two possible interpretations of 2.16. One is that Smith actually said that the impact is unprecedented (reporting). Another possibility is that Smith did not actually describe it as unprecedented, but may have argued instead that the impact is alarming or long-lasting and the writer therefore evaluates it as unprecedented. This clearly relates to issues of 'faithfulness' in reporting (Short et al. 2002). If the latter interpretation is correct, then 2.16 is non-reporting. If it is not, then 2.16 is reporting. In 2.17, due to the use of the 'action' verb (examined), this example is clearly a report of research, not of speech, and is thus non-reporting.

Examples like 2.18 instil scepticism regarding some aspects of Swales' model, at least in the current thesis. It renders Swales' view, that all non-integral citations with no reporting verb are non-reporting, incongruous. The first question to be asked in this example is about the role of the name of the cited source. The first, and most sensible interpretation is that the proposition is cited from Smith, thus reporting. The second, and less likely interpretation, is that only some part of this proposition is attributed to Smith, but the writer includes no boundary between averred and attributed information, if there is any. Therefore, I am inclined towards the first reading and see this example as reporting whereby the responsibility is shared (Groom 2000) or *reclaimed* by the writer (Hunston 1989, 2000). Along these lines, Pecorari (2013: 164) convincingly argues that "non-integral citation mitigates some of the force of the attribution; the writer presents the proposition as a fact which could simply be averred". This is true for Example 2.18. The writer presents the information as a fact which he can simply aver (see 'factive' propositions in Thompson and Ye's 1991; cf. Thompson's 2001 function Source as a function of non-integral citation). Yet, he draws on the source (Smith) for additional support.

Swales' distinction influenced many studies (e.g. Thompson and Ye 1991; Hunston 1993, 1995; Hyland 1999, 2000, 2002; Thompson and Tribble 2001; Ngah and Sze 1997; Borg 2000; Thompson 2001, 2005a; Charles 2006; Petrić 2007; Okamura 2007, 2008a, 2008b; Shoostari and Jalilifar 2010; Soler-Monreal and Gil-Salom 2011;

Jalilifar and Dabbi 2012; Manan and Noor 2015; Rabab'ah and Al-Marshadi 2013). Most of the aforementioned studies (e.g. Manan and Noor 2015; Thompson 2001, 2005a) were influenced by Thompson and Tribble's (2001) extension of Swales's model. Thompson and Tribble (2001) further sub-divided Swales' integral and non-integral citations as part of a framework relating citation forms to citation functions. This view resulted in the identification of three sub-categories of integral citations and another four of non-integral citations. Integral citations are either 'verb-controlling' (citation as the agent controlling a verb as in *Smith (2000) defines*), 'naming' (citation is a [part of] noun phrase) or 'non-citation' (the source is given but without the year of publication). Non-integral citations are either 'source' (attribution to a source of an idea), 'identification' (identifying an agent within the sentence), 'reference' (writer refers reader to another text, usually signalled by imperatives such as *see*) or 'origin' (originator of a concept or a product).

It is remarkable that the vast majority of these studies have linked citation form to function of, intention behind, citing (e.g. Thompson and Tribble 2001; Okamura 2007, 2008a, 2008b). Explicit recognition of the writer's intention and its relevance to forms is described in Okamura (2007). There are a number of issues associated with the identification of the writer's intentions (for detailed discussion of these see Sub-section 2.3.4). The aforementioned studies do not investigate the correlation between form and stance, particularly in non-integral forms with no reporting verbs (this aspect is analysed in Chapters 4 and 6, as part of implicit stance).

We now turn to studies of evaluation and stance in citations. These studies use resources from within evaluation or appraisal theories to understand how writers express stance towards cited authors or engage (i.e. interact) with their audience. Another strand of research looks further at how writers report the position of cited sources (e.g. Thompson and Ye 1991; Hyland 1999; Fardi et al. 2017). Stance and evaluation are used interchangeably to refer to the writer's assessment of cited information.

2.3.3 Research on stance, evaluation and appraisal

Having discussed the formal aspect of citation (in 2.3.2), I now turn to studies of 'evaluation' within and beyond citation. While the focus in this sub-section is on the literature on *evaluation in citation*, I shall start with a brief introduction about evaluation and the terms related to evaluative language. From this chapter onwards, an

important distinction will be made between 'writer-source evaluation' and 'writer-reader interaction' (or 'reader positioning', to use Hyland's 2005c and Lancaster's 2012 term).

The literature contains various approaches to the broader concept of evaluation. Evaluative language has been studied under different labels such as stance (e.g. Poppi 2004; Hyland 2005b, 2012, 2014; Charles 2006, 2007, Biber 2006; Englebretson 2007; Hunston 2007: Du Bois 2007; Coffin 2009; Thompson 2012; Lancaster 2012; Jalilifar and Dabbi 2012; Hyland and Jiang 2016; Kafes 2017), Appraisal (Martin and White 2005) and evaluation (e.g. Hunston 1989, 1993, 1994, 1995, 2000, 2004, 2011; Hunston and Thompson 2000; Thompson and Hunston 2000; Thompson and Zhou 2000; Römer 2008). Evaluative language has also been studied under other labels, which cover some but not all of the same ground, such as evidentiality (e.g. Chafe and Nichols 1991; Barton 1993; Aikhenvald 2004), point of view (e.g. Scollon and Scollon 1997), (authorial) voice (e.g. Hyland 2012; Thompson 2012; Dressen-Hammouda 2014), identity and metadiscourse (e.g. Hyland and Tse 2004; Hyland 2005c; Malmström 2007; Thompson 2012) and style (e.g. Johnstone 2009). The literature in this regard draws attention to the use of various linguistic resources such as hedges, boosters, (im)personal pronouns (e.g. Hyland 1994, 1998, 2005b, c) in addition to issues of alignment, solidarity, responsibility and commitment (e.g. Martin and White 2005; Du Bois 2007).

The literature has also investigated stance resources in various genres such as abstracts (e.g. Hyland and Tse 2004), Master's and PhD theses (e.g. Charles 2006; Thompson 2001, 2005a, b, 2009, 2012), undergraduate essays (e.g. Aull and Lancaster 2014), L2 writing (e.g. Hyland 2004) and research articles (Hyland 2012; Kafes 2017). Discussion of variations in methodology is beyond the focus of this thesis (for good summaries see Ochs 1989; Hunston 2011).

Of particular interest here are the notions of stance, evaluation and Appraisal.[1] These terms reflect variance in the use and methodologies of investigation into the broader phenomenon of evaluative language. Hunston (2011, Chapter 2) provides a compelling summary of the different approaches, and highlights points of consensus and contrast. I shall briefly introduce each term and then explore in greater detail as to how they relate to the current study.

Stance, a term first introduced and investigated by Barton (1993) and Biber and Finegan (1988, 1989), has been extensively researched in the literature (e.g. Du Bois 2007; Coffin 2009). The majority of these studies investigate the linguistic resources (also known as markers) which reflect the speaker's (Englebretson 2007) or writer's attitudes towards a proposition or an object. For example, Biber (1988, 2006) quantifies predefined lexical and grammatical markers of stance in different registers. Likewise, to some extent, Hyland (1999, 2000, 2009) quantifies predefined stance markers. However, Hyland (1999, 2000, 2001, 2005a, b and c) makes more explicit the view of stance as a social phenomenon whereby stance is a form of dialogue and engagement with the projected reader.

Various definitions have been proposed for the term stance. One commonly cited definition defines stance as "the lexical and grammatical expression of attitudes, feelings, judgements, or commitment concerning the propositional content of a message" (Biber and Finegan 1989: 93). Stance is also described as the "personal feelings, attitudes, value judgements, or assessments" of a speaker or a writer (Biber et al. 1999: 966). Both definitions strongly reflect the subjectivity of stance. In the last twenty years, stance has tended to appear as an element within 'metadiscourse', which Hyland and Tse (2004: 157) define as containing the "writer's stance towards … content or the reader". There is also the view of *evaluative stance* in Woodward-Kron (2002, 2003), Hood (2004), Englebretson (2007) and Lancaster (2012). In this view, evaluation forms part of stance itself. In other words, stance is presented as the umbrella term which covers evaluation, and other kinds such as epistemic stance. It should be noted that identifying evaluation in such research depends more on phraseology than individual lexical items. However, Thompson and Hunston (2000) include stance as one element within evaluative language. They define evaluation as "the broad cover term for the expression of the speaker or writer's attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about" (ibid: 5).

Taking a social view of stance, Du Bois (2007) conceives stance as a social tri-act, something which the stance-taker achieves in a specific context. In this view, stance is seen as a large unified phenomenon comprising various facets. Within stance, Du Bois includes different interrelated subsystems that together form his 'stance triangle' framework. His stance triangle framework contains vertices of evaluation, position and

alignment (see the following figure, taken from Du Bois 2007: 163, and included here as Figure 2.3).

Remarkable in this figure is the direction of each relation between the interlocutors and the object(s) being evaluated. The subject evaluates one or more object(s) or proposition(s), this act of evaluation in turn reflects the subject's position. Similarly, Hunston (2011: 12) asserts that evaluation "construes an ideology" that is shared between the text's writer and its potential reader. Du Bois states that evaluation and

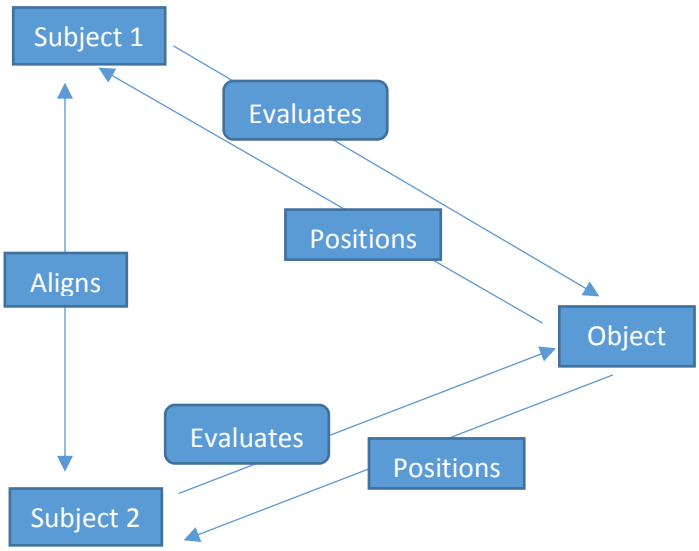

Figure 2.3: The stance triangle

positioning also highlight some form of alignment between the subjects, regardless of whether this alignment is convergent or divergent. Whether the evaluation of the object is initiated by the first or second subject, context shows how each subject positions themselves in relation to other subjects (or to the same subject, Du Bois calls this "self-positioning", p. 176), for example as either leading the evaluation or following or rejecting it. In sum, Du Bois's model (2007) provides an alternative view of stance which combines the various triplets of stance into one unified system. However, his model lacks discussion of the ways in which writers position their readers, as opposed to other subjects, and/or anticipate reader responses to current positions being established in the text.

Hunston (2011) notes that the term stance is used in two different ways in the literature. The first group of researchers identifies grammatical and lexical markers and then

quantifies the items' occurrences in the text, thus taking a quantitative approach (e.g. Conrad and Biber 2000; Biber 2006; Hyland 2005b). Englebretson (2007) and Du Bois (2007) belong to the second group which views stance as a social action. Englebretson's (2007) and Du Bois's (2007) approach stance as an action involving different steps in argumentation such as evaluation of the cited material and positioning of the speaker/writer. As a result, their approach enables "detailed exploration of individual instances" (Huston 2011: 23).

Taking a dialogic view of stance, White (2002, 2003) and Martin and White (2005), perceive all verbal communication as fulfilling a dialogic function. In this interpretation, speakers and writers engage with other subject positions and present their views, not only to evaluate an object but also to indicate their anticipation of reader responses. The dialogue is either 'expansive' (opening the space for alternative reader views) or 'contractive' (positioning the reader to accept the writer's view as the only possible position and thus closing the space for alternative views); see White (2002, 2003) for detailed discussion of dialogic engagement.

Similar views of language as a form of dialogue have been proposed in the literature using terms such as 'intersubjectivity' and 'engagement' (engagement here is used as in Hyland 2005b, 2005c; Hyland and Tse 2004; Hyland and Jiang 2016 to mean writer-reader, or speaker-listener interaction, following Johnstone 2009). Hyland (2005b), for example, quantifies features of stance (e.g. boosters, hedges and attitude markers) and engagement (e.g. 'directives', 'reader pronouns' and 'personal asides') in 240 research articles from eight disciplines. Under engagement, he investigates linguistic features which position the reader, reflecting the writer's awareness of the readers' expectations.

Martin and White (2005) develop the Appraisal theory within the Systemic Functional Linguistics (SFL) tradition based on analysis of newspaper discourse. In SFL, language is seen as a semiotic process in which writers select meanings to achieve three key 'meta-functions'. The first construes a world of experiences (and is known as the 'ideational' meta-function). The second is the 'interpersonal' meta-function and is realised via resources which exhibit the relationship between the 'interlocutors' (be it the speaker, writer and/or listener or reader). The third is 'textual', which functions to organise the discourse. See Halliday 1994, 2004; Butt et al. 2000; Eggins 2004 for more detail on each meta-function.

Appraisal, as a theory within SFL, investigates the resources of interpersonal meanings and is defined as "the semantic resources used to negotiate emotions, judgements, and valuations, alongside resources for amplifying and engaging with these evaluations" (Martin 2000: 145). Within Appraisal, Martin and White (2005: 35) identify three main systems which constitute the Appraisal resources, namely: 'Attitude' (which concerns the speaker/writer's feelings, judgement and evaluation and is realised through resources of 'affect', 'judgement' and 'appreciation', respectively), 'Engagement' (which concerns resources that position the speaker/writer with respect to his/her value position and also to that of the putative reader, such as 'proclaim', 'disclaim', 'entertain' and 'attribute') and 'Graduation' (which concerns adjusting the degree of evaluation, mainly divided into 'force' and 'focus'). Detailed discussion of the (sub)systems appears in Martin (1992), Martin and White (2005) and Martin et al. (2010).

At the core of the dialogic view is the Engagement system. Engagement looks into the ways in which the text's internal authorial voice is positioned in addition to the ways of positioning the reader's potential responses. Martin and White's (2005) view in this regard mainly derives from Bakhtin's (1981) and Voloshinov's (1995) notions of dialogism and heteroglossia which treat language, be it spoken or written, as dialogic and anticipatory of readers' responses.

Martin and White (2005) make an explicit comparison between their own Engagement model and Sinclair's distinction between averral and attribution (see Sub-section 2.3.1 above). They describe Sinclair's approach to attribution as "retrospective" (ibid: 135), meaning that it looks back into who says what in the text. Martin and White's Engagement model, on the other hand, pays less attention to the source of argument in the text. Martin and White are more concerned with "how the authorial voice [of/in the text] is positioning itself with respect to the anticipated reactions and responses of the audience which is being construed for the text" (ibid: 135). According to this view, the Engagement model has a "prospective or anticipatory orientation" (p. 135) in that writers construe for themselves some type of audience and that, depending on the way they present information, they anticipate and force the readers (or listeners) into the direction of either accepting, rejecting, or being undecided about, the proposed argument. Put differently, Sinclair regards attribution as a text-internal feature, whereas Martin and White see it as looking outside the text, particularly to the reader.

In contrast to Martin and White (2005), Sinclair demonstrates that identifying responsibility (who says what) can be, and usually is, important in the understanding and interpretation of averred and attributed propositions (see Sinclair 1986a, 1987; Hunston 1995, 2000). Drawing the line between what is averred and what is attributed in a text, or even a proposition, facilitates the understanding of how responsibility can be 'relegated' or 'reclaimed' (Hunston 2000; see also Thompson and Ye 1991; Groom 2000 on how responsibility can be shared in attribution). Martin and White's view above (which treats the source of arguments as "secondary" (2005: 135) to issues of how the arguments are presented to anticipate readers' responses) stands in sharp contrast to issues of averral and attribution (Sinclair 1987; Tadros 1993), faithfulness in reporting (Short et al. 2002) and writer and/or source evaluations (as in 'writer acts' and 'author acts' where the former signals the citer's evaluation of the source or cited material whereas the latter concerns the citee's evaluation of reported information). To illustrate the difference between Martin and White's Engagement system and Sinclair's approach to attribution, consider the following invented example (included as Example 2.19):

Example 2.19:

> Scientists speculate that there will be a dramatic increase in temperature in the Amazon Basin.

Martin and White perceive the above example as engaging the putative reader with what scientists say (or speculate). Put differently, the dialogic view within the Engagement system conceives Example 2.19 as the writer creating interaction between what scientists say and the rest of the world. The writer chooses to present the scientists' proposition as a speculation (here the issue of faithfulness in reporting comes into play as well; see Short et al. 2002), thus suggesting that scientists may be right or wrong. The use of the reporting verb *speculate* exemplifies distance as within the sub-system Attribute. The writer thus opens up the dialogue for the audience for alternative views (cf. the use of *point out* in Example 2.20 below).

Sinclair (1985, 1986a, 1987, 2004), on the other hand, argues that the attributed proposition *there will be a dramatic increase …* is embedded in an averred clause *Scientists speculate that* (see Sub-section 2.3.1). The writer's use of the verb *speculate* suggests that the scientists' proposition is not necessarily true and that we (the readers)

might expect the same writer to disagree with the attributed proposition as the text unfolds in full. In this view, Sinclair looks back into the origin of the argument, to identify who is responsible for it and also to see how the attributor presents it in the text.

Following Sinclair's steps, Hunston (1993, 1995, 2000, 2004) approaches evaluation (as in Example 2.19) with a particular focus on the interplay between text averrals and attributions. She is interested in the ways in which writers modify attributions in texts (e.g. by the use of some verbs or nouns of attribution; see Hunston 1994, 2000; also Thompson and Ye 1991; Thompson 1994, 1996). Particularly relevant to the above example is her evaluative function 'Status', which analyses discourse on a certain-uncertain scale, see below.

Hunston does not analyse examples such as 2.19 in terms of engagement between scientists and the world, as do Martin and White. Instead, she argues that the same sentence offers an evaluation of the attributed proposition *there will be …*. At the same time, she argues that this information comes from scientists, who are generally deemed reliable, so we might believe it. On the other hand, it is only that *Scientists speculate* (as opposed to *prove, evince, demonstrate*), therefore demonstrating uncertainty by the attributor for the attention of the audience.

Since 1989, Hunston has pioneered the research on evaluation in written discourse, producing novel research into evaluation functions (1989, 1993, 1994, 2000, 2007, 2011). She identifies three broad functions of evaluation, namely: 'Status' (making an object out of propositions alongside a certain-uncertain scale), 'Value' (evaluating objects and propositions on a good-bad scale) and 'Relevance' (how (un)important to the discussion are particular evaluations').[2] Hunston (2000: 6) notes that evaluation is specifically important 1) to express the writer's opinion, and thus reflect their ideology, 2) to construct and maintain relations between writers and readers, and 3) to organise the discourse.

Like Martin and White (2005), Hunston (1994, 2011) points out that evaluation is frequently "implicit" because academic writers quite often avoid attitudinal language (1994: 193). However, unlike Martin and White (2005: 61-8), Hunston does not distinguish levels of implicitness. She demonstrates, nonetheless, that quantitative

approaches to stance cannot capture implicit stance, which gives advantage to text-based studies of evaluation.

An extensive review of approaches to evaluative language covering a wide range of research into stance, Appraisal, evaluation and metadiscourse appears in Hunston (2011). Hunston discusses points of similarity and difference between the various approaches to the study of evaluative language, which can be summarised in the following points. In terms of variation, Hunston draws attention to studies which treat evaluation as an action in which the stance-taker engages (e.g. Englebretson 2007). Another realm of research investigated in Hunston (2011) perceives evaluation as a 'set of words and phrases' which express stance (e.g. Conrad and Biber 2000; Biber 2006; Hyland and Tse 2004). The third group treats evaluation as an interpersonal system in which writers choose from among the various resources to build relationships with their audience (as in the Appraisal system, see White 2003; Martin and White 2005). The final group, including Hunston (1989, 2000) and Hunston and Thompson (2000), investigates evaluation as a function of, or within discourse.

In terms of areas of agreement between the various approaches to evaluation, Hunston (2011) identifies six similarities. There is a general agreement that evaluative language 1) is subjective and intersubjective, 2) construes a shared ideology by the writer with their readers, 3) is realised by a wide range of lexical and grammatical items, 4) is contextual and cumulative, 5) involves an object or a target of evaluation, and 6) "permeates even the most objective of discourses" (ibid: 19). However, it could also be argued that the final point applies only to text-based studies, and not to those which quantify explicit markers of stance, such as stance adverbials (e.g. Biber 1988; Biber and Finegan 1989; Conrad and Biber 2000; Hunston 2007; Peacock 2015). In other words, while there is a relative agreement in the literature that evaluation can be highly implicit, quantitative studies which are corpus-based cannot account for implicit evaluation as context, phraseology and more importantly the analyst's interpretation play vital roles in decoding the implicit meanings in the text (see Section 4.2 in Chapter 4 for discussion of implicit stance).

The research on stance, evaluation and Appraisal discusses, to varying degrees, issues associated with attribution and reporting. Before exploring the literature on evaluative meanings within citation, I shall start with a brief discussion of certain aspects related

to citation from the literature. To start with, the extensive research on stance and/or evaluation has attended to issues of attribution (e.g. Hunston 1995, 2004; Coffin 2009), reporting (e.g. Thompson 1994, 1996; Caldas-Coulthard 1994) and discourse representation (e.g. Caldas-Coulthard 1994; Semino and Short 2004; McIntyre and Walker 2011).

In terms of attribution and reporting, it has long been recognised that reporting verbs convey different meanings and that the choice of the reporting verb has a strong effect on how the writer's view is presented and how the reader is positioned to respond to the cited work (e.g. Thompson and Ye 1991; Hyland 1999; 2002; Charles 2006; Bondi 2009; Diani 2009; Bloch 2010; Soler-Monreal and Gil-Salom 2011; Tafaroji and Boghayeri 2014; Jafarigohar and Mohammadkhani 2015; Fardi et al. 2017).

Thompson and Ye (1991) pioneered the strand of research on evaluative and denotational potential of reporting verbs. Their analysis resulted in the useful distinction between verbs which show the citer's evaluation ('writer acts') and those which reflect the citee's position ('author acts'). Following their lead, a number of studies have examined the use of reporting verbs in various contexts (e.g. Thompson 1994, 1996; Hyland 2000; Hunston 1993, 1994, 1995, 2000, 2004; Fardi et al. 2017).

As shown above, the system of Appraisal focuses on the internal voice of the text and how the information is presented to the reader. Within Appraisal, the Engagement system presents two forms of language: monogloss and heterogloss. In the former, there is "no recognition of dialogistic alternatives" (Martin and White 2005: 100) whereas in the latter, the writers present the information in a way which recognises the existence of alternatives. All forms of citation and attribution are heteroglossic in that they acknowledge the existence of "antecedent authorial voice" (Thompson 2012: 121), regardless of whether the writer agrees with that voice or not. Citation (and attribution in general) is analysed under the label Attribute in Martin and White. Under this heading, Martin and White include "formulations which [merely] disassociate the proposition from the text's internal authorial voice by attributing it to some external source" (2005: 111, emphasis added).

Martin and White (2005) make a distinction between the two sub-systems Proclaim: Endorse on the one hand and Attribute on the other. The key difference is that the former is contractive in that it closes the space for alternative views whereas Attribute

is expansive in that the writer clearly distances themselves from the proposition, attributing it to external source(s). Where the reader is positioned to accept a proposition, this is described as Proclaim, and is realised in the sub-systems 'Concur' (e.g. *of course, Admittedly … but*), 'Pronounce' (e.g. *Indeed*) and 'Endorse' (e.g. *X demonstrates*). Endorse is distinguished from 'Attribute' because Martin and White make it clear that attributions disassociate reported propositions. Where the space is open for either acceptance or rejection occur the expansive sub-systems Entertain (e.g. *perhaps*) and Attribute (comprising 'acknowledge' and 'distance'). By a way of comparison with Example 2.19 above (*Scientists speculate* …), consider the following invented example (2.20):

Example 2.20:

Smith (2000) points out that the destruction of Rome may have caused the war.

In this example, the argument in the reported clause *the destruction of Rome may have caused the war* originally appears in Smith, thus it is attributed. At the same time, the writer avers that *Smith (2000) points out that* and therefore the writer shares responsibility with the source. This is achieved by the use of the reporting verb *points out* which construes the citer's "commitment to the truth of the attributed statement" (Hunston 1995: 135). In other words, the writer of that statement evaluates as true or likely Smith's interpretation of the cause of the war, which may be *the destruction of Rome*.

In the evaluative sense portrayed by Hunston's work, the citer in the above example evaluates the attributed clause as true or very likely. Hunston's evaluative functions Value and Status play a key role in the analysis of such examples. The writer is committed to the truth or likelihood of the proposition even though the source (Smith) is not, because the attributed statement is hedged by the source (i.e. *may*). In other words, Smith presents *the destruction of Rome* as a possible (not definite) cause for the war. He thus acknowledges alternative views which might argue that the destruction of Rome may not have caused the war. The attributed statement is therefore heteroglossic, as it acknowledges (implicitly) different views.

In summary, it could be argued that the uncertainty associated with *may* is over-ridden by the contractive sense that the reporting verb *points out* conveys. The attributee (Smith) is praised for considering or proposing this possibility regarding the cause of

the war. If the above example was only averred (i.e. with no attribution), then it would be read as expansive since the use of *may* in that sense would have been an example of Entertain (allowing or bidding for alternative views). Since it occurs as the object of an *endorsed* attribution, however, it becomes contractive.[3] In other words, *may* seems to reflect uncertainty by the citee, but not for the citer who presents Smith's argument as true, or as the likely cause of the war.

Martin and White's Appraisal model, while usefully drawing on the resources of intersubjective positioning and dialogic engagement, does not deal with issues related to the interplay between averral and attribution in such instances and the consequence this has on the development of, and interpretation of, reported speech. Because evaluation is highly context-dependent and co-text plays a decisive role in deciding how resources from the various sub-systems act in a given example, it is possible to find some resources from Entertain (e.g. *may*, as in Example 2.20) acting contractively. This suggests that, contrary to how Martin and White (2005) would argue, the issues of responsibility, commitment and who says what are important, particularly in the analysis of academic discourse. See Example 2.21 below for further discussion of the interplay between averral and attribution and the importance of whose textual voice is being expressed in the text (the ambiguity of this example is discussed in Coffin 2009).

Martin and White's (2005) model has been highly influential in citation analysis research. A considerable part of the literature uses resources from within the Appraisal theory. For example, Coffin (2009) draws on and extends some aspects from the Appraisal theory to investigate citation practices in one PhD thesis in the field of Film Studies. She identifies four stance types (*acknowledge, distance, endorse* and *contest*) and some lexical items which realise the four stance types in the student's thesis. Stance in Coffin's study is perceived as a dialogic phenomenon for the expression of the writers' attitudes towards references and positioning of their readers.

Coffin (2009) analyses stance markers at the proposition-level with a clear emphasis on integral citations, possibly because there is a preference for such forms in the thesis she analysed. Coffin successfully draws attention to the ambiguity associated with certain formulations of non-integral referencing (ibid: 184) and quotes one example, represented here as Example 2.21:

Example 2.21:

> The 1930s was also a turning point for the Western – it saw the bifurcation of the genre into two sections with differing levels of popularity. The 'A' Western suffered a sudden and precipitous drop in its production, while the 'B' Western flourished and reached its peak of popularity, arguably providing the economic motive for the resurgence of the 'A' Western in the 1940's.[fn]

There are a number of interpretations of the source of *arguably*. One interpretation is that it is averred by the writer. Another interpretation is that it is attributed to the referenced author(s), referred to by the superscript number (quoted as [fn]). Coffin (2009) argues that the use of such formulations (i.e. non-integral references and what she calls 'assimilation'; that is, when writers reword sources' propositions via summarising or paraphrasing) obscures the heteroglossic nature of arguments, "creating a more (seemingly) monoglossic and therefore objective text" (p. 184).

In Example 2.21, it is unclear whether the writer's arguments are reinforced or advanced in the literature (Coffin 2009). Coffin goes on to argue that it is unclear from the above example whether it is the source who makes the argument *The 1930s was also a turning point for the Western* and whether it is the same source that questions (via *arguably*) that the resurgence of the "A" Western was caused by the flourishing of the "B" Western. Coffin is also undecided about whether both propositions are advanced in the cited reference and adopted by the writer or whether it is only the last position (*arguably*…). She contends that the rhetorical effect of this presentation is "to give both propositions a greater degree of objectivity, authority and weight" (ibid: 185). Following Coffin, I would argue that it is unclear from the above example which parts of the text are averred and which are attributed. The use of such formulations, I argue, enables student writers and published researchers to manipulate the resources of implicit stance and fuse their arguments with those of previous researchers. This creates the effect of making texts less subjective and thus giving prominence to the text's internal authorial voice.

Hunston (1995, 2011) investigates evaluation, particularly the function Status, in conjunction with attribution. In various publications (e.g. in 1995), she analyses the typical and non-typical meanings associated with four verbs of attribution (*acknowledge*, *insist*, *argue* and *claim*). The main argument in her study (also in later work, e.g. Hunston 2000, 2004) is that attribution is always evaluative and that verbs

of attribution construe multiple layers of evaluation (1995, 2000). She argues that an attributor can evaluate more than one entity or proposition in a single clause.

Hunston argues that some verbs (e.g. verbs of 'concession' and 'contradiction' such as *acknowledge* and *insist*, respectively) have certain "connotations" (1995: 156) which play a role in the interpretation of their discourse functions in context, even when the co-text gives no clue as to how the words are going to be read. In other words, an example such as *Johns insists that …* implies that the source (Johns) is unjustifiably assertive about a proposition that is unlikely to be true. Hunston (1995) draws a distinction between 'argumentative' and 'non-argumentative' reporting verbs and notes that with the former group (e.g. *argues, speculates*) there is a tendency for implicit evaluation, usually implying disagreement between the different parties involved. The disagreement is usually identified as the text unfolds (see the analysis of Beckerman and Pasek's (1995) paper in Chapter 6).

Hunston (2011) examines the linguistic resources associated with the analysis of Status (by analysis she means identification and labelling). She goes on to investigate the relationship between Status and attribution (see Section 3.3 in her text), pointing out that Status can be modified in a number of ways, one of which is attribution (see Hunston 1995, 2000, 2011). At the same time, attribution can reflect a dialogue between different statuses (Hunston 2011: 33-39). In simple terms, depending on the linguistic resources used in the modification of Status in attributed statements, the analyst should be able to identify whether the proposed information is attributed to the source, whether the source attributes or evaluates information from other sources and/or whether the writer reclaims responsibility for the attributed information (as in the Example 2.20, *Smith (2000) points out …*). To some extent, her research in this respect resonates with some studies of citation and reporting (e.g. Swales 1986, 1990; Thompson and Ye 1991; Hyland 1999, 2000; Thompson 2001), faithfulness in reporting (Short et al. 2002) and thought and discourse representation (Semino and Short 2004).

2.3.4 Research on citing behaviour: citer motivation, intention, reasons and rhetorical functions of citations

Under this heading, I will review the literature on citing behaviour and attempt to summarise different approaches to the study of citation function. It has been argued (Bornmann and Daniel 2008: 67), that "the development of citation classification schemes has not been a cumulative endeavour". Depending on the purpose and focus

of each study, the various schemes have produced different definitions and typologies of citation functions.

This sub-section reviews the literature on citation functions, motivations and intentions in relation to four questions, which are:

1- To what extent do researchers distinguish between citation functions on the one hand, and motivation, intention and reasons on the other? Is there a principled link between these terms or are they conflated?
2- What evidence do the researchers use to propose their categories of citation functions or motives?
3- What is the role played by form and stance in defining the functions?
4- To what extent do the various studies on citation function pay attention to evidence beyond the immediate sentence in which the citation occurs?

In the literature, the study of citation function initially appears as a subsidiary aspect of the study of citer behaviour (e.g. Broadus 1971; White and Wang 1997; Oppenheim and Smith 2001; Adel and Garretson 2006; Petrić 2007, 2012; Lee 2010; Mansourizadeh and Ahmad 2011; Harwood and Petrić 2012; Lynn 2014). Many studies in Information Science (IS) and Sociology of Science (SOS) have investigated citation function under the label 'citer motivation' (e.g. Garfield 1962, 1972; Hooten 1991; for extensive reviews of IS literature on citation analysis see Smith 1981; Small 1982; Cronin 1984; Pierce 1987; McCain and Turner 1989; Liu 1993; Baird and Oppenheim 1994; Case and Higgins 2000; Leydesdorff 1998; Ahmed et al. 2004; White 2004; Bornmann and Daniel 2008). From the 1970s onwards, there has been a considerable amount of criticism in linguistic research levelled against the views of IS and SOS research, particularly those which treat all citations as equal, such as quantitative measures of research contribution (see Section 2.2). The most common views in linguistic research into citation role in the citing text include Moravcsik and Murugesan (1975), Chubin and Moitra (1975), Murugesan and Moravcsik (1978) and Thompson and Tribble (2001). Since then, a growing body of the literature has focused on the 'rhetorical function' of citation from a purely linguistic perspective (e.g. Petrić 2007; Harwood 2009, 2010; Mastura and Azlan 2013; Lin et al. 2013; Gol et al. 2014; Shahri and Soghondikolaei 2015; Dontcheva-Navratilova 2015, 2016; Fardi et al. 2017).

In linguistic research, citation function has been studied under different labels such as citer 'behavior' (e.g. Shi 2008; Petrić and Harwood 2012), citer 'intention' for (non-)citing (e.g. Thompson 2001; Okamura 2007, 2008a, 2008b; Harwood 2009) and citer 'motivation' or 'reason' (e.g. Harwood 2008a, 2008b, 2010). Some of that research includes examination of student writing at different levels such as undergraduate essay writing (e.g. Lee 2010), postgraduate essay writing (e.g. Borg 2000), MA dissertations (e.g. Petrić 2007; Samraj 2013; Mastura and Azlan 2013) and PhD theses (e.g. Thompson 2001: 2005a, b, 2009; Thompson and Tribble 2001; Coffin 2009; Solar-Monreal and Gil-Salom 2011); as well as published writing (e.g. Campion 1997; Harwood 2009; Gol et al. 2014; Shahri and Soghondikolaei 2015; Dontcheva-Navratilova 2016).

In general, with respect to the first question posed above, the existing literature conflates function analysis of citation on the one hand with the writer's motivation, reason and intention analysis on the other hand. Citation function can be seen as a textual phenomenon while the latter three can be seen as human mental phenomena. There is a tendency in the literature however, to treat the four terms as synonymous. Studies which treat citation function as a textual phenomenon (whereby function is identified using textual cues, such as in Petrić 2007) or as a mental process (whereby function is related to citing writer's motives/intentions for (non-)citing, such as Harwood 2008a, 2008b, 2009, using what is known as "citer-oriented approach" Bornmann and Daniel 2008: 58) tend to equate citation function with writer intention or motive.

Harwood (2009: 516), for example, writes "Because an interview-based approach collects data about citing authors' purposes and intentions, functions can be identified regardless of whether there is any explicit indication of rhetorical purpose in the text". However, there are a number of issues with this argument. First of all, Harwood presents function as the inevitable consequence of the writer's identified intention. In other words, Harwood asks the writers about their intentions when citing and conflates their intentions with the citation functions, regardless of whether the writers are (un)successful in achieving their intended functions. There is a motivation or intention behind every citation which does not necessarily match the observed textual function. Second, Harwood makes a decision to exclude the role of textual cues and readers' perspectives in identifying functions. Another issue with Harwood's conclusion is that

it seems to conflict with his own assertion that there are reliability issues associated with interview-based informants responses such as "informants' recall error and lack of awareness" (2009: 500).

As noted in Section 2.2, only a few studies, such as Petrić and Harwood (2013), draw attention to a useful distinction between citation function (the rhetorical role of a citation) and motivation (reason for citing). Motivation, in that sense, is more "opaque" than function since it is not observable in the text and is determined by external factors outside of the citing text (ibid: 115; cf. Cronin 1981; MacRoberts and MacRoberts 1989, 1997; Kurtz et al. 2005). It could also be argued that even intention and reason for (non-)citing are not observable in the text. Only the citing author can account for his/her intentions or reasons for citing.

Although Petrić and Harwood (2013) acknowledge the distinction between citation function and motivation in theory, they do not distinguish it in practice. In their study of one student's self-reported citation functions, Petrić and Harwood (ibid) refer to the student writer's intended purposes as citation functions. They write that their respondent describes "what purposes she intended her citations to fulfil in her text, which we [Petrić and Harwood] refer to as citation functions" (ibid: 115). This clearly indicates that they conflate the writer's intentions with the citation functions. Along these lines, a number of studies discuss writer's intention when accounting for the rhetorical functions of citation (e.g. Petrić 2007; Harwood 2010; Harwood and Petrić 2012).

Conflating citation function with motivation and intention essentially means conflating the writer's perspective with that of the analyst and the reader. Texts, whether spoken or written, are produced with the putative audience in mind (Thompson 2001; Hyland 2005b). For example, academic papers in Harwood's case (2009) and student texts (Harwood 2010; Harwood and Petrić 2012; Petrić and Harwood 2013) are all produced for intended readership. Identifying the writer's motives or intentions is different from identifying citation functions. The reader and/or analyst are responsible for the textual functions of citations (their observed roles in texts), taking into account the available signals provided by the writer. Therefore, the reader and analyst perspectives are arguably more significant than that of the writer, in identifying the textual function.

Based on the evidence used to propose the function/motive categories (Question 2), the literature on the two subtly-variant fields (function and motivation) can be divided into interview-based and text/corpus-based studies of citation functions (Dontcheva-Navratilova 2015, 2016). It could be argued that interview-based studies of citation functions consider citation function as a mental process (e.g. Brooks 1985, 1986; Cozzens 1989; White and Wang 1997; Wang and White 1999). In other words, they perceive citing as a mental phenomenon and therefore they believe that the best way of arriving at the citer's intentions, motives and functions of citing is via the use of interviews (e.g. Harwood 2008a, 2008b, 2009, 2010; Harwood and Petrić 2012; Petrić and Harwood 2013), surveys (e.g. Vinkler 1987; Bonzi and Snyder 1991; Liu 1993b; Shadish et al. 1995) and/or questionnaires (e.g. Brooks 1984, in White 2004: 108). The text/corpus-based approach uses textual evidence in the immediate citing statement to arrive at the citation functions and also intentions (e.g. Thompson and Tribble 2001; Thompson 2001, 2005a, 2009, 2012; Petrić 2007; Samraj 2013; Dontcheva-Navratilova 2016).

The use of interviews is first documented in LIS studies of citer motivation. Brooks (1985, 1986; cited in Bornmann and Daniel 2008) was the first researcher to interview scholars about their particular motivations for some of their citations. Since then, a growing body of the literature in LIS and linguistics has adopted the interview-based approach in eliciting citer motives or intentions for (non-)citing. In the past decade or so, particular attention has been directed towards interviewing the citing authors to identify the citation functions (e.g. Harwood 2008a, 2008b, 2009, 2010; Harwood and Petrić 2012; Petrić and Harwood 2013).

Of particular interest here is the work of Harwood (2008a, 2008b, 2009, 2010), including recent work with Petrić (2012, 2013), which is based entirely on interviewing authors. As an example, Harwood (2009) identifies citation functions based on semi-structured interviews with twelve informants, obtaining twelve computer scientists' and sociologists' accounts of their citation functions in one of their research articles. The writers' responses were taken as evidence of citation function. Harwood's analysis produced eleven rhetorical functions (see Table 2.3) and resulted in the identification of evidence of intra- and inter-disciplinary commonalities and variations.

Harwood (2009) describes his functions as either enabling or allowing writers to achieve their intentions. As the definitions in Table 2.3 imply, Harwood takes it for granted that the reported intentions achieve the authors' purposes. He presents the functions as equivalent to intentions. Likewise, Petrić and Harwood (2013) arrive at their list of citation rhetorical functions by elicitation of one student writer's (Sofie) self-reported citation intentions.

Table 2.3: Harwood's (2009) taxonomy of citation functions, based on writers' perceptions of their motivation for citing (pp. 501-510).

Harwood's functions	Harwood's definition of each function
Signposting	Citations which 'direct readers to other sources'
Supporting	Citations which justify the writer's topic, method/methodology and/or their claims
Credit	Citations which acknowledge writer's debt to others
Position	Citations which show source(s) viewpoints
Engaging	Citations which reflect the writer's critical dialogue with sources
Building	Citations which show writer's use of source's methods or ideas
Tying	Citations which align writers with source's methods/methodology, positions and/or findings
Advertising	Citations which alert readers to previous work by the same writer or others
Future	Citations which establish future research plans
Competence	Citations which display the writer's expertise or ability to conduct research
Topical	Citations which show writer's research to be 'concerned with state-of-the-art issues'

Harwood refers to the interview-based approach for eliciting citation functions as "emic" (2009: 497). This means that he establishes the significance of citation function by interviewing writers about the intended functions of their citations. It is true that the 'emic' approach, as used by Harwood (2008a, 2008b, 2009, 2010; Harwood and Petrić 2012; Petrić and Harwood 2013), provides an insider perspective and thus better accounts for motives and intentions. However, it does not necessarily account for citation functions.

In order to understand Harwood's emic approach, looking at emic research more generally is useful. For this particular purpose, I draw on the work of Silverman (2010). One of the cases he reports is of a student researcher (Tippi) who studies "living and coping in a community of elderly people" (Silverman 2010: 44). In that study, the researcher (Tippi) interviews a random sample of eight elderly people to see how they feel about living in their community. Tippi's aim was to clarify "the basic meaning of living [in that community]" (cited in Silverman 2010: 44). She asked the elderly people about their daily routines, attitudes to neighbourhood, services, environment and relatives as well as their past experiences of living in the same society to see how they would compare early-life experiences with current (at the time of interviewing) experiences.

There are various advantages associated with such an approach. In the case of Tippi's study interviewing the locals reveals community-specific meanings and experiences to which the researcher does not have access without the locals' assistance. Tippi can only observe the social practices in that community but, in order to understand and explain their culture-dependent meaning(s), she needs an insider's perspective.

Harwood uses the term emic to refer to the general practice of eliciting citation functions using interviews. He criticises text-based approaches for not adopting such a method. In citation analysis, the advantage of an interview-based approach is that it can account for the writer's intentions or motives for not citing.

There are, of course, various differences between Tippi's and Harwood's use of interviews. In her use of interviews, Tippi purposively samples (see Silverman 2010: 141-3) the community to account for the elderly people's views of living and coping in the society. It would not be appropriate in her case to interview young locals, for example, or to observe the locals' practices and reflect on them. Without an insider's

perspective, Tippi would not be able to reflect on the hidden meanings or ideologies of the inhabitants. Therefore, her sampling and method are arguably the best available for her research.

Harwood's case is slightly different. While he attempts to categorise the citation functions by interviewing the citers, he arrives at a conclusion that linguistic signals are insufficient to indicate either the purpose of citation or its effect. However, there are various ways to arrive at citation functions without interviewing the citing authors. Context plays an important role in the identification of citation functions (Hellqvist 2009). In fact, many typologies have used textual evidence to propose citation functions (e.g. Thompson and Tribble 2001; Thompson 2001, 2005a, b, 2009, 2012; Petrić 2007). It could be argued that there are other emic approaches which use textual cues to produce and explain citation functions.

There are, of course, various limitations to interview-based emic approaches. For example, Tippi found that her respondents recounted similar narratives. Silverman (2010: 45) draws attention to this point and thoughtfully asks: "How far is it appropriate to think that people attach a single meaning to their experiences?" In other words, it is suggested that in Tippi's study, there might be "multiple meanings of living in the community" (ibid: 45) which could be represented by other informants or even expressed by the same informants but only to social carers or between themselves and not to the interviewer. Therefore, Silverman points out that a major methodological issue with interviews is that they seem to be "actively constructed narratives" (see also Riessman 2008; Gubrium and Holstein 2009, cited in Silverman 2010: 45) rather than providing direct access to insider experiences.

The same criticism applies in Harwood's case, I would argue. While Harwood's interview-based approach is useful in identifying citer intentions based on their comments, it remains subject to criticism, as Harwood himself acknowledges. For example, there are issues which might risk or affect informants' responses such as their "recall error", "lack of awareness" (ibid: 500), "self-serving" responses (where some respondents may avoid functions which are less "savoury"), ability to "introspect" (2008a: 1011, 2009: 517) or unwillingness to cooperate. Further, there are occasions when some authors failed to explain their motivations or felt "at a loss to explain why" they cited a source in a certain way (Harwood 2008a: 1010). Citing authors'

interpretations of their intended functions are subject to change over time or between different co-authors. There is a possibility that if a paper is written by two authors, then they may offer two interpretations of the motivation for some citations. It is even possible that a single author may change his interpretation over time. This risk is still present in approaches that are based on the analyst's perspective. Yet, the analyst usually follows explicit linguistic markers in the text which makes their interpretation less likely to change over time, as opposed to methods which rely on the writer's ability to recall their intended functions. Taking these issues into account, it seems reasonable to argue that there are multiple meanings (or intended functions) of citing, represented by the citers' narratives and what they say to the interviewer and to each other.

In general, discourse-based interviewing is a powerful instrument typically used to elicit the authors' perceived motives for their citations (e.g. Hyland's 2003 study of self-citations and Harwood's 2008, 2009, 2010 studies of motivations and functions) or of other writers (e.g. Cronin 2005; Harwood 2008a, 2010). However, extensive review of the literature reveals that interviews are more suitable for analysing citation motivation from the perspective of the citing author, but they are less powerful in explaining the functions of citation. Within the citing behaviour literature, motivations for and intentions behind citing reflect the writer's choice and reason for using citations. Citation functions on the other hand, can be more confidently identified in the immediate co-text within, and beyond, the citing statement.

The text/corpus-based approach takes into account textual signals which are available at the disposal of the analyst at the time of analysing citation functions. While the writer's motive for, or intention behind, certain choices in their text may only be accurately identified by the writer himself/herself, the role of the analyst is to identify how and for what textually-observed, not intended, purpose a citation is used, whilst concurrently taking into account various aspects such as linguistic cues and context. To demonstrate the difference, it could be argued that one motive for citing is to satisfy an examiner's or publisher's expectations, whereas the function might be to Support an ongoing argument in the text or Signpost the reader to an external source.

Text-based studies identify explicit textual cues (e.g. Berkenkotter and Huckin 1995; Hyland 2000; Thompson and Tribble 2001, Thompson 2000, 2001, 2005a, 2009, 2012; Charles 2006; Coffin 2009; Mansourizadeh and Ahmad 2011; Samraj 2013), the lack

of which may create problems for the analyst, and the citer, in extracting or recalling the intended functions. It is true that expert writers may occasionally, via skilful manipulation of linguistic resources, conceal the textual functions. However, careful analysis of a citation in its context of operation reveals how the citation functions in respect of the surrounding context. The text-based approach is therefore important since citation function analysis can exhibit inconsistency between what the writer and the present context have to offer for the intended and observed functions, respectively.

One text-based approach that is frequently adopted in the literature (e.g. in Gol et al. 2014; Dontcheva-Navratilova 2016) is that of Petrić (2007). Petrić examines the function of citations in sixteen high- and low-rated MA theses in gender studies. She adopts Thompson's (2001) functional classification of citations and further adds more rhetorical functions. Her analysis resulted in the identification of nine rhetorical functions (see Table 2.4, overleaf).

In her study, Petrić makes explicit that "The main criteria for assigning a citation to a particular category were linguistic cues in the citing sentence signalling the writer's intention in relation to the cited source" (2007: 242). There are a number of limitations to her approach. For example, she draws on 'linguistic cues' to derive the 'writer's intention' behind citing sources. As noted above by Harwood (2008a, 2009), the linguistic cues are insufficient to reflect the writer's intentions. They can however, enable the analyst to deduce the textual function of citation, as Samraj (2013) notes. Linguistic markers could alternatively be seen as indicators of the role of citations in the text from the perspective of the analyst, not from that of the writer.

In terms of Question 3, some of the literature using text-based evidence of citation function draws on resources from within citation form (mainly Swales 1990) and stance in reporting (mainly Appraisal theory according to Martin and White 2005, and Thompson and Ye's 1991 classification of reporting verbs). Thompson and Tribble (2001) establish the relationship between citation form and function (see also Thompson 2001: 104 on the role of "formal features" in revealing the "intention of the writer"). Thompson and Tribble (2001) produce a revised typology of citations which builds on and extends Swales' (1990) formal distinction between integral and non-integral citations. They add a functional aspect to Swales' formal typology of citations, combining formal and functional criteria to account for various functions observed in

integral and non-integral choices (Table 2.5 summarises their form-function categories). Thompson and Tribble's typology has profoundly influenced various investigations of citation function (e.g. Okamura 2007; Lancaster 2012; Manan and Noor 2015; Shahri and Soghondikolaei 2015; Ali 2016; Dontcheva-Navratilova 2016).

Table 2.4: Petrić's (2007) taxonomy of citation functions, based on the analysis of textual signals at the level of the citing statements (pp. 243-246).

Citation functions	Petrić's definition of each function
Exemplification	Citations which provide information on the source
Statement of use	Citations which state which works are used by the writer
Attribution	Citations which attribute an activity or information to a source
Further reference	Citations providing further information on an issue
Establishing links between sources	To indicate differences between sources, showing writer's ability to identify 'controversial issues'
Application	Citations which show writer's use of source's ideas, methods or terminology
Evaluation	Citations which show writer's evaluation of source(s)
Comparison of one's findings with others	Citations which indicate similarity or variation between writer's and source's findings/interpretations
Other	Citations in which "the relationship between the citing statement and the citation is obscure" (p. 246)

In various publications, Paul Thompson (e.g. 1999, 2000, 2001, 2005a, 2005b, 2009, 2012) has pioneered the analysis of citation practices of PhD theses, particularly on the correlation between citation form and function. He examines different aspects related to the practice of citation, analysing postgraduate students' citation practices to demonstrate how, via the choice of whether to cite and how to cite, students can give prominence to research over authors (2001, 2005a), identify responsibility for propositions (2005b), and achieve various functions.

Table 2.5: Form-function categories from Thompson and Tribble's (2001) framework (these categories also appear in Thompson 2001)

Citation form	Citation functions	Brief explanation
Integral citations	Verb controlling	citation acting as an agent controlling a verb
	Naming	citation as a noun phrase or part of a noun phrase
	Non-citation	citation without 'a year reference', p. 96
Non-integral citations	Source	attribute propositions to other sources
	Identification	identifying an agent within the sentence
	Reference	writer referring readers to external sources
	Origin	indicate the originator of concept, method or product

Moving on to the role of stance in defining citation functions, to the best of my knowledge, very few studies (e.g. Coffin 2009 and Ali 2016) have used some resources from stance models (e.g. Appraisal theory) in addition to identifying citation functions (see also Thompson's 2001 use of 'factive' citations from Thompson and Ye 1991). The vast majority of citation studies treat stance and function as two variant textual phenomena and thus approach them in isolation from one another. One of the studies

which uses resources from within stance and also identifies rhetorical functions of citation is Coffin's (2009), in her analysis of citation practices in one PhD thesis in the field of film studies.

Using a modified version of the Appraisal theory, Coffin identifies four stance types (*acknowledge, distance, endorse* and *contest*) and nine rhetorical functions. However, Coffin does not align the stance and function categories. In other words, the identified functions are not identifiable from the four stance types, therefore she does not discuss the relationship between stance and function in citation analysis. She associates all the rhetorical functions in her typology to acknowledging citations (this is where the student writer attributes information or ideas to another source) because *acknowledge* is the most common stance type in the thesis she analysed. Yet, the reason for not aligning function and stance categories is not articulated by Coffin. This is probably because the identification of citation functions has not been specifically linked to the role of citations in the ongoing argument.

While the stance categories all reflect the writer's evaluation of the source, the identified functions operate in different directions inside and outside the citing text. For example, one function in Coffin's typology is "Comparing and Contrasting Theoretical Views" (Coffin 2009: 189). This function seems to reflect the student writer's intention/reason for citing but it does not necessarily reflect the role of the citation with respect to the ongoing argument (e.g. whether it Supports or opposes the writer's argument). It is common for a writer to identify contrast between two cited studies, and at the same time, use one of these studies to support their argument. Coffin's model captures examples of the latter use (e.g. "Strengthening Position By Referencing Evidence", p. 191) but does not account for variation between text-internal and text-external functions, those which directly relate to the writer's argument as opposed to those which compare or contrast cited sources. In this sense, function analysis in her study seems to capture the various reasons for citing, which is different to the citation role in the "successful construction of an argument" (ibid: 188). In other words, she identifies functions which are common for the writer's purposes in such a genre but the identification has not been formulated in a way that could be aligned to stance identification; specifically to find out how each citation acts in the citing text with respect to the citer's argument and/or finding.

Coffin's framework was recently used in conjunction with Thompson and Tribble's functional framework of (non-)integral citations (Kafes 2017) in a cross-cultural study of citation practices. Kafes investigates citation practices in novice Turkish academic writers and expert native English speaking academics in the field of applied linguistics. His study reveals that differences outweigh similarities between the two language groups due to individual differences in terms of language proficiency and level of subject-knowledge.

Ali (2016) analyses stance in citation following Thompson and Ye's (1991) classification of reporting verbs and later refinements by Hyland (1999, 2002). He specifically examines 'research acts' (e.g. *find, examine*), 'cognition acts' (e.g. *note, realise*) and 'discourse acts' (e.g. *discuss, propose*). In terms of functions, he uses Thompson and Tribble's framework to identify errors made by Sudanese writers and finds that the Sudanese writers make various errors in citations, unlike their British counterparts, in the PhD theses they produced. Like Coffin (2009) and also Kafes (2017), Ali (2016) does not align stance and function categories in order to capture the variant ways in which citations are evaluated and utilised in the citing text. He also analyses integral citations only while the model he uses (Thompson and Tribble 2001) accounts for integral and non-integral citations.

Generally speaking, the citation function literature identifies evaluation as one function of citations. Evidence of this may be found in various studies, such as Petrić (2007; and more recent studies such as Petrić and Harwood (2013) and Gol et al. (2014). Petrić specifically includes evaluation as one citation function. Similarly, Petrić and Harwood (2013) identify the functions of "Agreement" (p. 115) together with other functions such as "Support" (p. 114). The former function shows the writer's agreement with the information they cite whereas the latter enables the writer to justify their claims. This clearly indicates that evaluation and function are conflated and gives rise to a logical question. That is: if both functions of Agreement and Support are identified, will the citation be treated as multi-functional? And if yes, what are the criteria for identifying the functions? Alternatively, one might reasonably ask: why are evaluation and function not treated as two separate, yet related, phenomena and therefore aligned? It seems reasonable to argue that Petrić and Harwood's list does not form a binary system. That is to suggest that their taxonomy does not seem to reflect a principled decision as to which functions can be grouped together and on what basis. The same observation

applies to Harwood's functions in Table 2.3, particularly Engaging and Supporting on the one hand, with Advertising and Future on the other, and the functions of 'statement of use' and 'evaluation' in Petrić (2007). In this sense, treating evaluation as a function of citation seems to be an oversimplification of the role of citation in text. Citers draw on various choices when evaluating sources (e.g. positively, negatively or neutrally) to achieve certain functions. Functions could usefully be aligned with stance types to show how the citations are evaluated and for what textual (as opposed to intended) functions.

Having discussed the citation function literature in terms of the extent to which researchers distinguish functions from intentions, the evidence used to propose the categories and the role of form and stance, it is useful thereafter to discuss the role of the citation context in devising the various functions (Question 4). In general, the study of citation functions has been restricted to the level of the citing statement (e.g. Petrić 2007; Coffin 2009; Snodgrass 2011; Dontcheva-Navratilova 2016; Ali 2016) even though part of that literature identifies multi-functional citations (e.g. Harwood 2009; Petrić 2007; Samraj 2013). For example, Petrić identifies "multiple functions" (2007: 247), yet only uses "linguistic cues in the citing sentence" (ibid: 242). Snodgrass (2011: 13) found citations with "more than two functions", and yet "made a judgement call regarding which to use". This is because annotation in her study was made at the sentence level. Snodgrass concludes that "sentence level analysis does not necessarily line up with the broader communicative context" (2011: 49) drawing on Hellqvist's (2009) contention that citation function analysis in the humanities should exceed the sentence level for proper identification of the function(s).

Of particular relevance here is Hunston's (1993) work on evaluation. Hunston (1993: 131) argues that stance can operate in blocks. Although the term 'block' has not been specifically used in her study, she demonstrates that evaluation of opposed claims could be carried in consecutive sentences following the citing statement. She identifies various moves through which this evaluation can be achieved. It seems reasonable to argue therefore that stance beyond clause-level can be considered as an indicator of the role of the citation in the block. Put differently, if a citation is acknowledged at the clause-level but critiqued at the block-level, then the criticism could be taken into consideration in function analysis, since the observed pattern would be to *acknowledge*, in critical dialogue, the cited information. Function analysis therefore complements clause-level analysis of stance.

Only a number of researchers have investigated the role of citation beyond the proposition level. Thompson, for example, is one of very few researchers to acknowledge the role of citations beyond the level of the proposition (2005b, 2012). In so doing, he combines formal criteria, such as the presence/absence of resources of averral and attribution (following Sinclair 1988; Tadros 1993) which operate at the proposition-level, together with metadiscoursal features which operate beyond the proposition-level (following Hyland 2005b). This consideration of the position of the citation beyond the proposition-level assists interpretations of how citations operate in students' writing, at a higher discourse level (2005b, 2012). At this level, Thompson contends that writers can achieve and "maintain the dominance of authorial voice" (2005a: 321). He calls for sensitising the students to the formal features of citations, and also to the implications of various choices. As far as the discourse level is concerned, Thompson highlights that sensitising the students to their choices should enable students to position themselves in relation to their immediate audience as well as to the broader disciplinary community.

The role of context has for long been a key concern in automatic classifications of citation functions (e.g. O'Connor 1983 and Bertin et al. 2016). The general conclusion from this research is that the citation context plays a decisive role in the identification of citation functions. Quantitative measures and methods of identifying the functions are insufficient. Therefore, various contexts of analysis have been proposed; for example, to include one or two statements preceding or following the citation (e.g. O'Connor 1982), three sentences surrounding the citation (Schneider 2006) and n-grams (Sula and Miller 2014; Bertin et al. 2016). However, as Sula and Miller (2014) note, most of the proposed schemes lack explicit operational application by virtue of employing overt linguistic markers to enable automated analysis. Automated classification schemes tend to investigate large corpora and obtain quantitative results on the occurrence and frequency of certain linguistic markers. The problem with such large corpora (e.g. in 75,000 articles corpus analysed in Bertin et al. 2016) is that they leave little or no space for the investigation of context (Handford 2010). This is probably due to the "decontextualized" nature of corpora (Hunston 2002: 110, cited in Handford 2010: 256). In-depth context analysis is therefore important in the investigation of citation functions.

Bertin et al. (2016) use a combined approach, implementing n-grams and sentiment analyses at the level of the citing sentence, which they perceive as the context, to classify citation functions based on full-text analyses of 75,000 research articles published in seven journals. The authors conclude that limiting the citation context to the citing statement level results in many linguistic cues beyond the level of the statement being left unattended. Therefore, the authors propose, for future work, to model texts as sets of sentences instead of word sequences (i.e. n-grams), arguing that it is better motivated from the linguistic point of view. They admit that their quantitative methodology does not reflect the "complexity of the phenomenon related to in-text citations" (2016: 1431); an issue that they aim to tackle in future research via the implementation of an automatic annotation of citational contexts by more linguistically-motivated qualitative approaches.

The conclusion from the previous research is that context is indispensable in the identification of citation functions. The context boundary will remain a matter of dispute, however. Combining different levels of analysis may mitigate the associated effects of limiting context boundaries. The analyst will always play the decisive role in coding, verifying citer choices and stance(s) towards the cited materials and more importantly, the citation functions. It is possible to teach computers to divide citations into pre-identified contexts (e.g. paragraph in which citations occur). However, it is not possible to teach computers to decode the implicit evaluation, and therefore citation functions.

2.4 Conclusion

This chapter has introduced the concept of citation and some related terms such as attribution and reporting. It has also reviewed items from the literature on the central influences in citation research (e.g. form, stance and function of citation). Extensive research has been conducted on citation analysis and this chapter has served to provide a theoretical orientation towards the proposed model. LIS research has mainly studied citations as a gateway into improving retrieval performances and measurement of scientific growth. Within LIS, a sizable body of the literature has investigated the citer behaviour. The vast majority of this research focuses on motivation and reason for citing. A concomitant line of research in linguistics has also investigated citer behaviour. The findings from the studies in the two disciplines suggest a concern with

citation as a mental process and thus citations are seen as academic practices which reflect the writer's intentions when reviewing items of previous research.

The literature lacks a theory for citation analysis that takes into account the complex nature of citations and mainly their roles in context. This thesis aims to tackle this issue. The next chapter outlines the proposed framework, drawing on resources from some existing typologies. However, the proposed model extends and to a certain extent modifies certain aspects of the previous models. In other words, Chapter 3 shows how the choices and categories in this thesis differ to some extent from those in the adopted models. It also describes the data and important notions in the proposed model, such as the citation block.

Notes

1- When capitalised, Appraisal refers to the system as introduced in Martin and White (2005). All the sub-systems from Appraisal (e.g. Engagement, Attribute, Entertain) will be also capitalised. In Chapter 3, I shall distinguish between Engagement as in Martin and White's (2005) sense and as a function of citation, as in Harwood's (2009) sense.

2- See Hunston 2011: Chapters 2 and 3 for a detailed overview of evaluation functions and their relation to attribution.

3- In Martin and White (2005), Attribute includes instances which present an external voice, to which the internal authorial voice is either neutral or undecided. Thus, Example 2.20 above (*Smith (2000) points out*) is studied under the sub-system Proclaim: Endorse in Martin and White's (2005) model.

Chapter Three

Methods and the Proposed Model

3.1 Introduction

This chapter provides an overview of corpus collection methods and sets out the proposed model. The proposed model builds on and extends some of the linguistic approaches which appear in Chapter 2. The next section (3.2) briefly introduces the Interdisciplinary Research Discourse (IDRD) project (Thompson et al. 2017), noting the contribution of that particular project to the collection and sampling of papers in the current study. I have used resources originally compiled as part of the IDRD project (Thompson et al. 2017) in classifying and tracing diachronic and synchronic variation in the selected papers from within the journal Global Environmental Change (hereafter GEC). Section (3.3) gives a brief account of data collection and sampling methods. Section 3.4 then explains the process of identifying citations and suggests how and why it is different from other methods reported in the literature (e.g. Hyland 2000; Petrić 2007). The last section (3.5) answers questions regarding how the model was developed and how the various form, stance and function categories have been identified and aligned in a way to enable multi-level analyses of each citation at clause- and block-levels. This section is further divided into two sub-sections: the first sub-section (3.5.1) introduces the model categories while the second sub-section (3.5.2) introduces the new analytical unit, the citation block, which is used to identify citation functions and citation relations beyond the proposition-level.

In brief, this chapter describes the data and the model. It also introduces all levels of analysis and the resultant categories. This includes a definition of the levels, categories and citation relations identified at each level with diagrammatic representation of the various relations. In addition, the way in which citations have been identified is also briefly introduced in this chapter. In general, this chapter shows how the model has been reformulated and how it differs from the previous models on citation form, stance and function which are summarised in Chapter 2.

3.2 The IDRD project

As noted in Chapter 1, this thesis is indebted to the IDRD project, which was funded by the Economic and Social Research Council (ESRC). During the early stages of

corpus compilation, this thesis benefitted from the use of resources originally compiled as part of the IDRD project (which commenced in 2015; papers from the project were published in 2017 by Thompson et al. and Murakami et al.). The purpose of the IDRD project was to investigate interdisciplinary research discourse. This thesis also investigates interdisciplinary discourse but, unlike the project, this thesis attempts to account for citation practices in the journal GEC.

The IDRD project compiled a corpus comprising 11,201 full-length papers in eleven journals on environmental studies (Thompson et al. 2017). Thompson et al. further investigate texts in one well-established interdisciplinary journal, GEC. It is important to note that the IDRD project does not investigate citation practices. Instead, it uses factor analysis from Biber's (1988) Multi-Dimensional Analysis (MDA) model to identify dimensions of variation among academic papers in the field of environmental science. The dimensions are then used to compare the eleven journals, each representing a sub-corpus. The main purpose is to use MDA "to ascertain the degrees of linguistic variation between and within the eleven journals" (Thompson et al. 2017: 2). In order to explore this variation, the authors further investigate 673 papers from GEC from the period 1990-2010.

GEC is identified by the publishers (Elsevier UK, cited in Thompson et al. 2017: 6) as a successful journal that has maintained "its broad, interdisciplinary appeal" for more than two decades (since 1990). GEC interprets global environmental change as "the outcome of processes that are manifest in localities, but with consequences at multiple spatial, temporal and socio-political scales" (GEC website, retrieved 5 May 2018). GEC specifically welcomes papers with a "social science component", covering a wide range of issues related to the drivers and effects of environmental change and also social processes in addressing environment-related issues (ibid). GEC papers have a wide coverage of environment-related issues which include, but are not limited to: the climate, food systems, land use and cover, water resources, oceans and urban areas. GEC targets a wide audience such as policy makers, researchers in various disciplines and organisation managers (GEC website, retrieved 5 May 2018).

The IDRD project establishes that GEC is composed of constellations of papers, each with a particular dimensional profile. The profiles are developed from quantification and analysis of the occurrence and frequency of linguistic features from within six

dimensions from Biber's (1988) MDA model (e.g. formal versus informal features and explicit versus implicit argumentation; see Table 3.1 for a brief account of the dimensions and some associated features; for further detail on the dimensions and associated linguistic features, see Appendix 2 in Thompson et al. 2017).

Table 3.1: The MDA dimensions (from Biber 1988) with some associated features

The six dimensions	Some associated features
1- system-oriented vs action oriented	Present tense verbs, determiners, abstract nouns
2- explicit vs implicit argumentation	Use or no use of models, stance adverbs, conjuncts
3- informality/spoken-ness	Personal pronouns, contractions
4- conceptual discourse	Attributive and topical adjectives, connecting phrases
5- text-focused vs site-focused	Features of reported discourse, e.g. place nouns
6- non-research vs research worlds	Nominalisations

The examination of the dimensional profiles of the papers in GEC resulted in the identification of six clusters of papers, later called "text constellations" (Thompson et al. 2017: 13). Each of the six distinct constellations contains papers which share sets of features and thus represents a distinctive approach to the problem of global environmental change (Hunston and Murakami 2015).[1] Prior to the identification of the constellations, the papers were classified based on their types (e.g. *policy, empirical, research agenda/research framework* and *other*). The categorisation of the constellations came as a result of a second attempt to categorise and group papers which have commonalities in their approach to environment-related issues.

The identification of the dimension features allowed the authors to investigate the extent of variation between the constellations in GEC. Most of the variation was observed between constellations 1 (Quantification and measuring) and 5 (Personal

voices). Papers in constellation 1 scored low on dimensions 1, 2, 3 and 5, and high on dimension 6 (Thompson et al. 2017). The authors attribute these scores to concerns with actions rather than with systems, implicit argumentation, non-research world, space and place and the construal of human societies as "abstractions defined by environment-related activity" (ibid: 19). These features are usually identified in scientific research articles. Therefore, this constellation was taken to represent the scientific component of the GEC corpus.

In contrast, papers in constellation 5 scored high on dimensions 1, 2, 3 and 5 and low on dimensions 4 and 6, revealing opposing results to those of constellation 1. This suggests a focus on systems, explicit argumentation, features associated with informal language and the research world. As the title of constellation 5 suggests (Personal voices), papers in this constellation engage "with a number of voices" (ibid: 20), dealing with human perspectives on environmental issues. It was found that papers in this constellation often articulate an "antagonistic stance towards a purely 'scientific' approach to environment research (Thompson et al. 2017: 18). It was also found that some authors in constellation 1 are pessimistic about future environmental changes. However, papers in constellation 1 make specific reference to specific research locations and their anticipated changes are usually linked to the geographic location of study. Papers in constellation 5, on the contrary, engage with scientific positions on global environmental issues, including the empirical research reported on in constellation 1 papers. The main points of divergence between constellations 1 and 5 are summarised in Table 3.2 (overleaf).

3.3 The sample

As discussed above, the IDRD project reveals the highest variation between constellations 1 and 5, whereby constellation 1 represents the 'science-like' component of the corpus and constellation 5 represents the 'social science-like' component. Based on the results of the IDRD project, I expected to also find divergence in citation practices between the two constellations, both qualitatively and quantitatively. It is also possible that the two constellations will have either diachronically diverged or converged in their practices, which will be determined from the citation practices, which are elaborated on in Chapter 4 and quantitatively compared in Chapter 5. I also expected the citation practices to reflect the broader disciplinary communities from

which they come, since GEC is an interdisciplinary journal which receives contributions from different disciplines.

Table 3.2: Summary of differences between constellations 1 and 5 (as in Thompson et al. 2017)

Points of variation	***Constellation 1***	***Constellation 5***
Research focus	Space and place and Specific sites of human and policy interaction	Text and system
	Non-research world	Research world
	Physical world	Abstract world (world of ideas)
Argumentation	Implicit and less concern to explain steps of argumentation	Explicit, usually more signalling of the steps involved in argumentation
Approach	Quantitative data and construe human societies as abstractions defined by human-related activities "de-humanised approach" (p. 19)	Papers in this constellation deal with human perspectives on the environment and social perspectives of science
Stance towards other approaches/prediction	Papers attach value judgements to predictions about environment-related issues	Engagement with research voices and antagonistic stance towards other approaches
Overall dimensional scoring	Scores low on dimensions 1, 2, 3 and 5 and high on 6	Scores high on dimensions 1, 2, 3 and 5 and low on 4 & 6

Twenty GEC papers were chosen for analysis in this thesis, containing a total of 1186 citations. The corpus was divided into four sub-corpora, each sub-corpus consisting of five papers. The first sub-corpus relates to the early period (1990-95) of constellation 1. The second sub-corpus samples the later period 2008-10 of constellation 1. The third and fourth sub-corpora sample the early and recent periods of constellation 5,

respectively. This decision was influenced by the resulting number of papers in each sub-corpus. Constellation 1 contains 118 papers whereas constellation 5 contains only 35 papers altogether. Neither myself, nor Thompson et al. (2017), had control over the distribution of papers within the focal time periods and constellations and as a result, the papers were unevenly distributed, as is to be expected. The early period of constellation 5 only contains five papers; therefore, it was decided to include only five papers from each of the other sub-corpora to ensure a standardised sampling of the four sub-corpora. This also enables a small-scale synchronic and diachronic comparison of citation practices and the occurrence and combination of certain form, stance and function categories in the four data sets.

The literature shows that researchers on some occasions choose between "balance and representativeness" (Hunston 2008: 166) when considering the size of the different corpus components (i.e. sub-corpora). In this study, I chose to evenly balance the size of the four samples to enable a comparison of the findings between the four sub-corpora. As noted above, the small number of papers does not allow for generalisation nor did the analysis aim to represent GEC as a whole. Instead, I aimed to reflect tendencies in the selected papers, and give an indication of citation practices in the respective time periods and constellations. I also aimed to discuss the results of the current thesis in relation to those reported in the IDRD project, specifically considering variation in citation practices between the two constellations.

Diachronic convergence or divergence between the two constellations can be examined via an analytical comparison between the two time periods in each constellation. The first time period includes papers between 1990 and 1995 while the second period comprises papers between 2008 and 2010. In other words, I sampled the four sub-corpora not only to observe synchronic variation between the selected papers from the two constellations, but also to see if the papers in each constellation show convergence or divergence in terms of their citation practices over time. The quantitative results are presented in Chapter 5 and discussed with reference to the IDRD project in Chapter 7.

The analysis of citation practices in this thesis is mainly qualitative. This thesis takes an inductive approach and I originally identified and analysed citations to test and revise the adopted models on citation form, stance and function. Therefore, the ultimate focus was not on corpus size or the ability to generalise the findings related to

interdisciplinary discourse or academic writing.[2] Instead, I sampled the corpus with the aim of showing how the authors of the selected papers from the two constellations and time periods use citations and for what specific functions. In other words, the findings should not, and were not intended to, be generalised to the other constellations or to the remaining 653 GEC papers included in the IDRD project. The findings were used to revise the adopted models and to propose a new model that takes various aspects of analysis into account (see Section 3.5 on framework development). Briefly, the resulting model is a triad system which aligns form, stance and function categories to enable analyses of citations at various levels.

Bryman (1988) asserts that qualitative research should follow a theoretical, not statistical, logic and that the findings should be generalisable to "theoretical propositions" (p. 90) rather than to wider populations or universes. This link between sampling and theory was later taken up by Mason (1996) who argues that sampling in qualitative research should be linked to the researcher's theoretical position, their research questions and most importantly the "explanation or account which you [they] are developing" (pp. 93-4, emphasis mine). Sampling in qualitative research should therefore be "theoretically grounded" (Silverman 2010: 143). Along these lines, Hunston (2008) contends that the purpose of research plays an important role in its corpus design and that any corpus "cannot be judged except in the context of its purpose" (ibid: 155). For these reasons, and because of the fine-grained nature of analysis, the current thesis collects a small number of research articles in order to investigate citation practices in the 1186 blocks. In other words, I use the results of a corpus study (Thompson et al. 2017) in order to carry out a methodical collection of texts which is then used to textually analyse citations in contexts.

A small specialised corpus which allows for "qualitative, contextually-informed analyses than those carried out using general corpora" (Flowerdew 2004: 18; Lee 2008) is suitable for investigating citation practices of a particular specialised genre. I decided, therefore, that the corpus size in the current study was sufficient to analyse and compare citation practices in the papers from the two constellations and time periods. The ultimate focus was on the number of citation contexts to be examined, not on the number of papers. The 1186 citations produce an equal number of citation contexts (blocks) of operation. Analysis of the blocks is qualitatively dense, suggesting that this number of citations-in-blocks is sufficient for the current thesis. Further, in order to

minimise the risk of idiosyncratic citation practices, I compiled the corpus to equally represent the four sub-corpora and to provide the widest coverage of synchronic and diachronic variation between papers from the two most diverse constellations.

Extensive research has investigated the role of context in large corpus analysis (e.g. Widdowson 1998, 2000; Hunston 2002, 2008; Lee 2008; Handford 2010). The overall conclusion is that, at some point, a researcher might have to choose between context and corpus size depending on their research purpose(s) and question(s). There is evidence in the literature that large corpora do not allow the researcher to say much about context (e.g. Handford 2010). Some researchers have even described corpus data as not authentic because it is mostly taken out of context (Widdowson 1998, 2000; as cited in Handford 2010: 259). The GEC corpus can be seen as a specialised genre of interdisciplinary research articles on environment-related issues. My initial plan was to include as many papers as possible for citation analysis. However, due to space and time restrictions and the in-depth nature of the analysis, I decided to include twenty papers to be scrutinised for citation analysis. The twenty papers contain 1186 citations, each of which is analysed in its identified context of operation: the block (see Sub-section 3.5.2).

As noted in chapter 1, the objective of this research was to explore citation contexts. I made the decision to explore citations in their contexts when Hunston (2015, personal communication) drew my attention to a publisher's (Elsevier UK) concern regarding the lack of research into citation value, in addition to personal scepticism about the sufficiency of citation counts (e.g. h-indices) in reflecting the importance of scientific contribution and academic impact. The purpose of this research is not to investigate citation value, but rather to propose a model as a response to the limitations and inconsistencies of some existing models in their investigations of citation practices. The proposed model aims to better account for how citations act (function), and are acted upon (stance), in the citing texts. It uses the selected papers to scrutinise citation practices in GEC as a successful interdisciplinary journal.

3.4 Identifying citations

The method used in this study omits certain types of attribution; in particular, cases of anaphoric reference to a citation without repetition of the source's name and date of publication. Williams (2010: 618) calls this "implicit attribution" (IA). Consider the

following example taken from Williams (2010: 618), included here as Example 3.1 (the sentences have been numbered for ease of reference):

Example 3.1:

> [1] To [author James Q.] Wilson, the erosion of marriage is eating away at society's foundations. [2] More than anything else, it is responsible for the poor, troubled and wretched corners of America. (from Patricia Cohen, review of J.Q. Wilson, *The marriage problem. New York Times Book Review* 4/21/02)

Williams (2010) argues that the first sentence contains a citation to Wilson, even though the date is not present at the surface level of S1. He also argues that S2 exemplifies IA since the attribution is correctly inferred, even though "no source authorship is actually asserted in the sentence" (p. 618). Williams also argues that the writer intends the readers to construe S2 as a continuation of Wilson's opinion and therefore both statements "exhibit proper source attribution" (p. 618). A clearer example of IA is when the writer provides the source's name and date of publication in the first statement, and use a third-person pronoun to refer to the same writer in the second statement (e.g. *Smith (2000) argues that .... He also argues*). In this example, the source's name and date of publication appear as an integral component of the first statement, which is counted in the current thesis. However, the second statement, starting with *He also argues,* exemplifies IA in Williams' terms and like S2 in Example 3.1 is not included in the analysis.

The identification of citations is not as simple as might be expected and different researchers take different views in terms of what should and should not be included in citation analysis. For example, some studies (e.g. Hyland 2000; Petrić 2007; Dontcheva-Navratilova 2016) identify and count citations regardless of their "surface realizations" (Dontcheva-Navratilova 2016: 60). In other words, the presence of the author's name, whether or not it is followed by the year of publication of the cited study, is included for analysis, as in the study by Williams (2010), and only footnotes and endnotes are excluded from the analysis. This differs from the identification process used in the current study. Guided by the objectives of my research, I identify and count only the citations which contain authors' names and the year of publication and/or those which are included in the respective reference list. These are the occurrences which contribute to the calculation of citation counts and which have been used as indicators of research impact. The proposed model can also be applied to instances of implicit

attribution, as one can also identify stance in instances of IA (e.g. *He also argues*) and also the citation function, as long as the broader context is given.

Once the citations were identified, I adopted an 'inductive discourse analysis' approach as introduced and defined by Barton (2002; and later adopted in Samraj 2013). In this approach, the analyst repeatedly reads and analyses the data until patterns emerge. Citations were analysed and the model was concomitantly revised to categorise and code the citations. Each of the 1186 citations was analysed at the clause- and block-levels to identify form, stance and function categories. I manually coded and numbered the citations for ease of reference. The initial analysis took place in 2015 and lasted for about 3 months. The model was under development at that stage. Later, once the patterns started to emerge, particularly the conventional and unconventional combinations of stance and function categories (see Section 4.3 for examples), I re-analysed the citations and coded them in terms of the revised form, stance and function categories.

Because of the in-depth analysis of citations at various levels, it was not possible to recruit a second inter-rater to check the results of the first and second analyses. Therefore, I repeated the analysis for a third time to maximise accuracy. As an intra-rater reliability measure, the analysis of the citations in the third round confirmed the results of the second. The accuracy rate between the third and second analyses was 97%, which I deemed satisfactory. The reason behind this high rate is that the analysis depends on recorded textual cues in the identification and coding of the categories. I then used the coded categories of form, stance and function to devise the conventional and unconventional combinations. However, it is important to note that in a few doubtful cases I discussed the analysis with a second rater who also specialises in Applied Linguistics to ensure that the coding was accurate. These accounted for only a few cases and the discussion confirmed the original coding.

3.5 Framework development

This section gives a simplified overview of the proposed model. In general, the current thesis investigates citations in terms of the writer-author evaluation and the writer-reader interaction. Following White (2003) and Martin and White (2005), the current thesis uses resources from the system of Engagement: Attribute. It also adopts Bakhtin's (1986) view of texts as being involved in dialogue with other texts (cited in

Thompson 2001: 48). In terms of citation, this thesis investigates how writers evaluate and use other texts to achieve certain purposes and also to position their readers in terms of the cited information. This section sets the scene for detailed analysis of citations in Chapter 4. The purpose here is to introduce, with brief examples, the various categories which represent the various relationships identified in the texts. These categories will be revisited with further examples from the data in Chapter 4 to demonstrate how the categories relate to each other.

The current study combines three main relationships that have been identified in the analysis of citations, as summarised in Figure 3.1. As noted in Chapters 1 and 2, these relations seem to be examined in isolation of each other in the literature, particularly stance and function. However, I argue that in order to understand the roles of citation in text, one must consider these relations as they are indicative of the ways in which citations act, and are acted upon, in the citing text. In presenting these relations, this thesis modifies certain aspects of some existing models to suit the needs of this study, and more importantly, to make these relations comparable.

In Figure 3.1, the solid lines represent what is termed here the 'citing-cited relation': the stance taken by the citing writer towards the cited author. Stance is analysed at the clause-level of the citing proposition. There are four stance types, as adopted from Martin and White (2005) and a compelling summary in Coffin (2009); these are: *endorse*, *contest*, *acknowledge* and *distance*. Each of these types is defined in Sub-section 3.5.1 below.

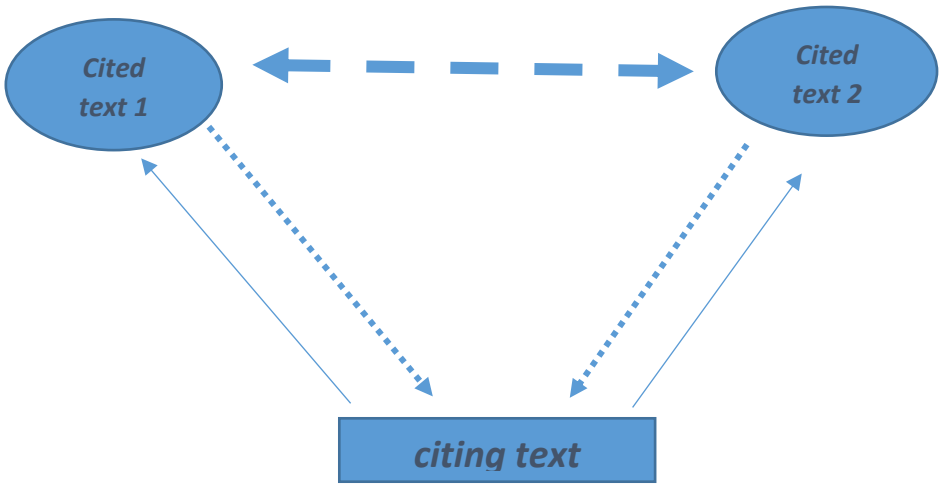

Figure 3.1: The relationships between the citing and cited texts

The dotted lines represent the extreme opposite of how citations are acted upon in the citing text. In this relation, the focus is on how citations act in the text, otherwise known as citation function or the 'cited-citing relation'. This relation is also initiated by the citing writer but, unlike stance, it describes the way(s) in which the writer employs the cited information to function in their text. This relation is analysed at a new level, called the 'citation block' (see Sub-section 3.5.2). Four functions have been chosen to represent this relation: Support, Engage, Signpost and Position.

Each of these functions corresponds to one stance type identified above in the same order of presentation. For example, *endorse* as a stance type corresponds to the function Support and every other stance category corresponds to one of the functions. This correspondence is termed here as the 'conventional' combination. However, the conventional combinations are not the only ones that have been identified. Unconventional combinations include, for example, the *endorse* stance type being used with the function Engage. See Sub-sections 4.3.1 and 4.3.2 for examples of both conventional and unconventional combinations. In this thesis, form and stance categories are identified at the proposition level of the citation. The citation functions are identified at the block-level and then aligned with the identified form and stance categories (examples are presented in the next chapter). In general, the combination of form, stance and function categories resulted in the identification of sixteen citation patterns, four of which are conventional and twelve are unconventional. These combinations have been identified in integral and non-integral forms.

Another relation identified also at the block-level is the 'cited-cited relation' and is represented by the dashed line in the above figure. This relation has been studied under another label, namely 'author act' (Thompson and Ye 1991; also referenced in Thompson 2001). In its simplest form, it represents the links that are drawn by the writer between two or more external sources. This can be achieved by the use of certain reporting verbs such as *refute* and *accept* (for further detail see Thompson and Ye 1991; G. Thompson 1994; Fardi et al. 2017) or certain structures which show (dis)agreement between external sources (see Section 4.4 for examples).

This relation includes the links drawn originally in the cited sources as well as those established by the citing writer and is realised in two broad categories, what I term 'citation-support' and 'citation-engagement'. The former is related to agreement

between sources while the latter relates to disagreement (see Section 4.4 for examples). It is important to note that this relation is distinguished from citation function in the citing text since the cited-cited relation relates to the links drawn between the sources. Function analysis, on the other hand, specifically relates each citation to the citing text irrespective of whether that citation shows agreement or disagreement with another cited text. In this sense, function analysis investigates the role played by every individual citation in the citing block.

3.5.1 Introducing the model categories

As noted in Chapter 1, the proposed model is a response to the limitations of the individual models of citation form, stance and function (some of these limitations appear in Chapter 2, Sections 2.3.2, 2.3.3 and 2.3.4, respectively). To the best of my knowledge, no study has attempted to align categories of the above individual models in order to devise a model that can be more comprehensively reflective of citation practices and their roles in text. Therefore, the present study addresses this gap by proposing a model that can be replicated in other research contexts. In studying citation form, stance and function, the current thesis revised the adopted models. This revision has resulted in the identification of ten categories, see Figure 3.2 (overleaf).

The proposed model is both deductive and inductive. It is deductive in that it applies previously discussed models of form, stance and function (mainly Swales 1990 on form; White 2003, Martin and White 2005, Coffin 2009 on stance and Harwood 2009 on function) as the basis for citation analysis. On the other hand, it is inductive because it uses authentic data to revise and modify certain aspects within the models (and their categories) in order to make them comparable and facilitate simultaneous use. The use of authentic data to reflect on existing models and to produce a revised taxonomy that takes into account the limitations of each individual model and integrates various levels of analyses should make the proposed model exhaustive in terms of the various levels of analysis and perspectives.

The bulk of this research revolves around three main aspects in the identification and analysis of citations. The proposed model for analysing citation forms a triad system in which each citation is analysed and coded in terms of form, stance and function. The following three tables (Tables 3.3, 3.4 and 3.5) summarise and define the main analytical categories.

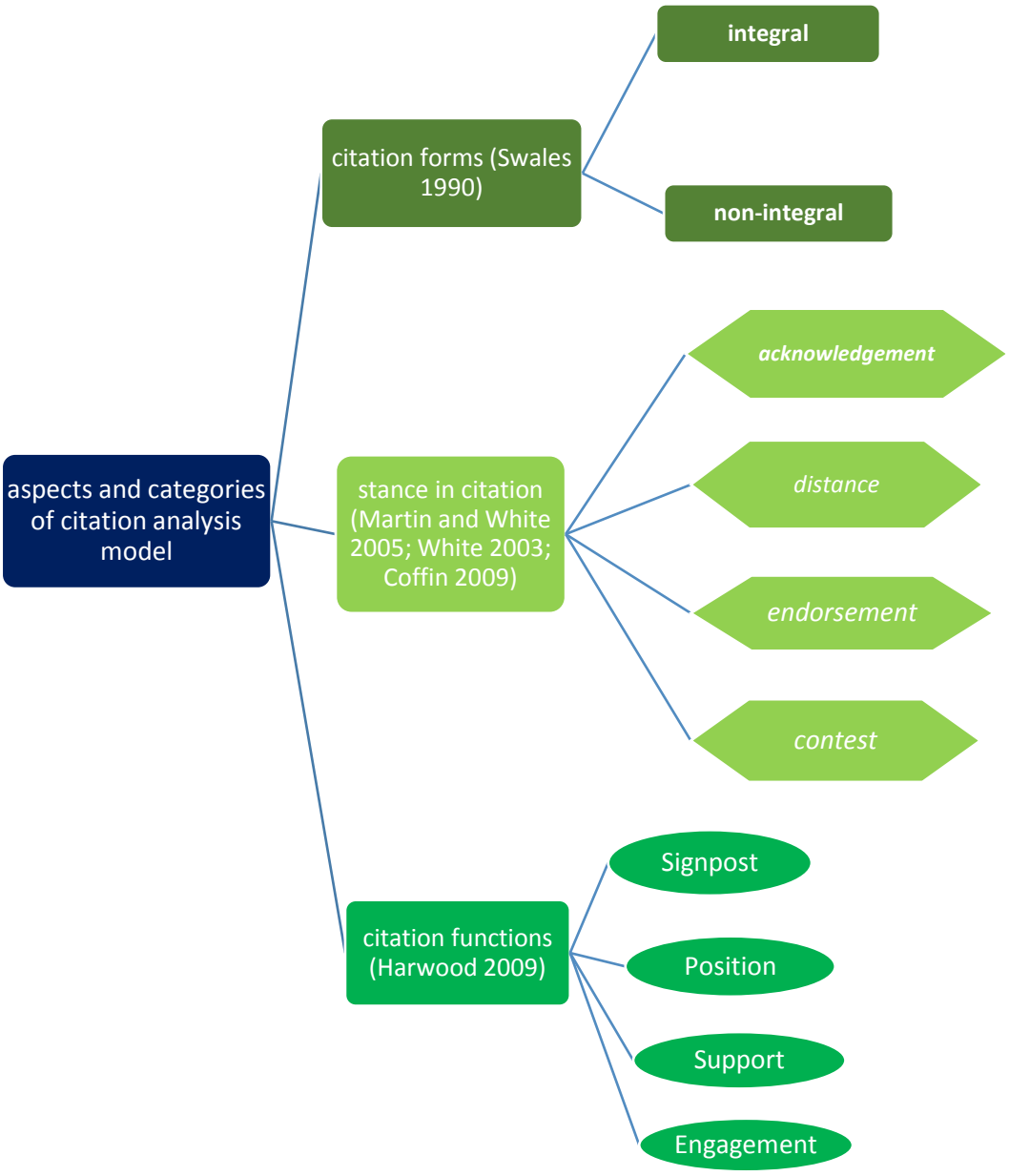

Figure 3.2: A summary of the analytical framework

Table. 3.3: Citation forms (Swales 1990)

Citation form	Definition
Integral citations	They form part of the citing statement, usually in Theme position as the subject or the participant, or as a passive agent
Non-integral citations	The author's name plays no grammatical role in this type. It is usually added in parentheses or footnotes.

Table. 3.4: Citation stance types (Martin and White 2005; Coffin 2009)

Stance types	Definition
endorse	Positive evaluation of cited materials, can be explicit or implicit. Writer shares responsibility and commitment with cited author(s)
contest	Negative evaluation of cited materials, can be explicit or implicit. Writer clearly does not share responsibility with cited author(s), includes total or partial criticism of cited information
acknowledge	Neutral reporting, no positive or negative evaluation of cited information
distance	Writer withdraws from cited proposition, no commitment or responsibility is shared with cited authors, and no negative evaluation of source's argument

Table. 3.5: Citation functions (despite methodological variation, the terms are derived from Harwood 2009)

Function types	Definition
Support	Citation is employed in support of the writer's argument/finding/main idea in the block (e.g. when a source *endorses* a citation beyond clause-level and uses it as the basis for their argument)
Engage*	Writer is involved in critical dialog with cited information or author** Engagement can be with some or all of the cited information
Signpost	Citation acts merely to guide readers to external sources
Position	Citation shows a source's position or viewpoint of a particular topic/argument, no commitment or shared responsibility on the part of the writer

* Engage/Engagement (with capital E) is used as a function of citation and is derived from Harwood's (2009) model.

** This function is different from *contest* in two ways. First, it operates at a broader level, the block. Second, while it could be argued that on such occasions the writer *contests* the author at the block-level, this function differs in the direction of operation (cited-citing) from *contest*. In other words, a cited source cannot criticise a citing source, at least in the citing block because the original text precedes its citing text. The citer can employ such citations to show that they are involved in critical dialogue with the source's information, arguments or findings. Further, *contest* as a stance type signals the citing-cited relation. That is, how the citing writer evaluates the cited author or their propositions at the clause-level. Engage, as a function, however, exemplifies the cited-citing relation. Like other function types, it shows how the cited information is employed by the writer to act in the citing block, to show cited material/argument as being inconsistent with, or different from, the citer's argument.

The distinction between citations, based on their forms, is adopted from Swales (1990). According to Swales, citations can be divided into integral and non-integral based on their syntactic form and position. This division is based on whether they form part of the citing proposition (integral) or play no grammatical role in the citing proposition

(non-integral). Non-integral citations normally take the form of references in parentheses or footnotes.

Swales' distinction is formal and gives no information about meaning. Nevertheless, it provides a sensible starting point in the identification of citations and one advantage is that there is unlikely to be any dispute over the identification of each category in this aspect (see Chapter 2, Sub-section 2.3.2.2). A considerable part of the literature is focused on integral citations. Some examples include Thompson and Ye (1991), Hyland (1999, 2000), Soler-Monreal and Gil-Salom (2011) and Ali (2016). Some of the above studies confine analysis of citations to integral forms focusing mainly on reporting verbs and structures. One of the key distinctions in meaning in citation practice is the choice of reporting verbs and tense. This has been extensively discussed in the literature. However, this thesis argues that similar distinctions also exist in non-integral citation practice. The choice of tense and reporting structure as well as presentation of the cited information can be indicative of stance and function of citation. The following are two examples of integral and non-integral citations:

Example 3.2:

> First estimates for global resource use **were presented by** Schandl and Eisenmenger (2006)… [2010_Schandl and West]

Example 3.3:

> Starting with two **influential** reports launched by the World Resources Institute in 1997 and 2000 (Adriaanse et al., 1997; Matthews et al., 2000) which compared material flows for five OECD economies, the methodological standardization ….. [2010_Schandl and West]

Example 3.2 contains an integral citation. The source appears as a passive agent in the phrase *were presented by*. Thus, the authors' names play a grammatical role in the construction of the passive voice sentence. In 3.3, however, the same authors made a choice to include the authors' names in parentheses. The writers chose to introduce the reports in the citing sentence, but also to non-integrally cite the sources. Thompson (2001) demonstrates that this has the effect of foregrounding previous research more than the researcher's contribution. In both examples, the writers chose to foreground research (*estimates* in Example 3.2 and *reports* in Example 3.3) but the difference is

that the authors' names constitute one part of the citing proposition in 3.2 and play no role in the construction of 3.3.

There seems to be a tendency in the literature to underrate non-integral forms as they are often treated as carrying no stance and to be used mainly to *acknowledge* external sources. However, this thesis provides evidence (e.g. *influential reports* in Example 3.3) that this view is an oversimplification and that non-integral forms are used in more sophisticated ways to express different stances.

The analysis of stance benefits from resources originally compiled in White (2003), Martin and White (2005) and Coffin (2009). Coffin (2009) lists four stance types in citation, which are listed in Table 3.2 and diagrammatically represented in Figure 3.3. This aspect of analysis relates to the writer-author (or citing-cited) relation. Stance in citation is equivalent to the citing writers' attitudes, comments, commitment and positions towards the cited propositions or authors.

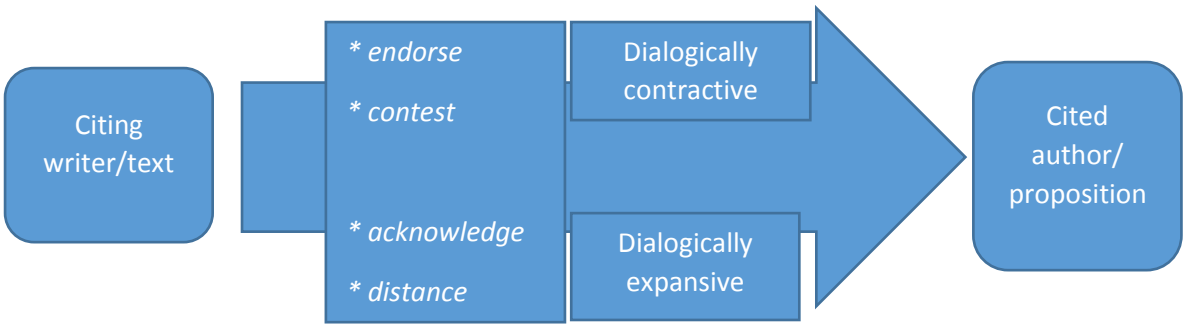

Figure 3.3: Stance towards cited information at the clause-level

As can be seen from Figure 3.3, this relation is concerned with the writer-author evaluation on the one hand and the writer-reader interaction on the other. Thompson (2001) analyses PhD theses and argues that texts are dialogic in nature reflecting interaction between the writers and their intended audience. Likewise, I argue that stance in citation reflects the relationship that is established by the writer towards the cited authors and the intended readers. The four stance types in the above figure indicate how the cited information is presented by the writer for their readers. In terms of the writer-reader interaction (following White 2003 and Martin and White 2005), the identified stance type indicates how the cited information is evaluated and dialogically positioned for the attention of the reader. The first two stance categories, namely, *endorse* and *contest*, are dialogically contractive in the sense that they close the space for alternative voices or stances (White 2003) by the reader. In being positive or

negative towards a cited piece of information, for example, by presenting that information as either factual or false, writers force their audiences in the direction of either accepting or rejecting the cited information, thus leaving no other sensible choice for the readers. *Endorsement* relates to positive evaluation of cited authors or information whereas *contest* is associated with negativity towards authors or cited information.

Endorsement can be explicit or implicit. Explicit *endorsement* uses explicit markers of positive evaluation such as some reporting verbs (e.g. *demonstrate*, *show*, *point out*) and structures (e.g. *as X found, as X argues*). Other markers include stance nouns (e.g. *disclosure, success*), adjectives (e.g. *influential*, *seminal*) and adverbs (e.g. *successfully*, *convincingly*, *compellingly*). Also, positive evaluation includes devices which show commitment to cited information as in boosters which "emphasize certainty or close dialogue" such as *indeed, in fact, definitely* and *it is clear that* (Hyland 2005c: 49; these resources appear under the label 'Proclaim' in Martin and White 2005). Martin and White (2005: 121) include under Proclaim formulations which "act to limit the scope of dialogistic alternatives in the ongoing colloquy". Figure 3.4 gives a diagrammatic representation.

Implicit *endorsement* does not use the above markers but it is recognised by the way the information is presented (see Chapter 4, Section 4.2). For example, the information is implicitly *endorsed* when the writer presents cited information as factual, true, reliable or coming from a trustworthy author or institution. In the next chapter, I discuss instances when *endorsement* is implicit.

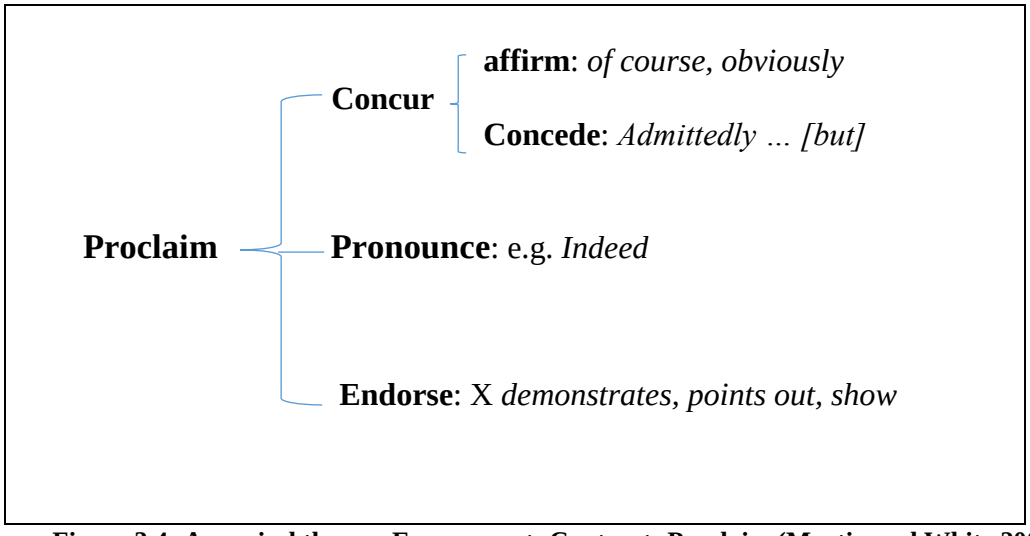

Figure 3.4: Appraisal theory: Engagement: Contract: Proclaim (Martin and White 2005)

Like the other stance types, *endorsement* of sources can be realised in integral and non-integral citations. Example 3.3 above shows *endorsement* in non-integral citations. Describing the reports as 'influential' makes it clear that the writers positively evaluate the cited source's work, at least at the clause-level. It should be noted at this stage, as a reminder, that form and stance are analysed at the clause-level. Stance beyond the level of the clause is taken into account in function analysis. The reason for this choice, as noted in different places in the current thesis, is to compare the results of clause- and block-levels and to identify the resulting patterns. Integral *endorsement* of the source is found in the following example:

Example 3.4:

> Fairhead and Leach's **pioneering** work on Guinea, West Africa (1995) provides a critique of these …. [2008_Dull]

The authors' names appear as part of noun phrase in subject position in 3.4. The writer is clearly *endorsing* the authors' work (as *pioneering*). These examples show that *endorsement* is commonly identified in both forms of citation. The same applies to the other stance types. Expressing explicit stance is common in academic writing in both citation forms. The difference is in the way the writers choose to present the information: that is, whether they give prominence to the source (integrally) or their work, findings or arguments (usually non-integrally).

Contest is usually marked at the clause-level, via the use of many resources to express negativity towards authors or cited information such as certain reporting verbs (e.g. *overlook, fail, lack*), or evaluative lexis (whether nouns, adverbs or adjectives as in *confusion, alarmingly* and *false*). That is not to suggest that implicit *contest* is unachievable. It rather means that, unlike implicit *endorsement*, it is uncommon in the data of the current thesis. It seems that some writers prefer to *acknowledge* sources at the clause-level and criticise (Engage with) them at the block-level (see Chapter 6 for examples). Returning to *contest*, the following example gives an account of how instances of *contest* are identified in the data. Again, it should be noted that *contest* occurs in integral and non-integral forms.

Example 3.5:

> This suggests that **rather than** treat the scientific community's priority setting as sovereign (**as** SBC **tends** to[fn]), an important role for the interpretative social sciences is to explore the debates and arguments within the scientific community … [1995_Shackley and Skodvin]

The writers in this example are clearly critical of SBC's approach. They call for an alternative approach to interpretative social sciences. This call is signalled by *rather than* and also the structure *as SBC tends to*. The footnote in this example contains the full reference entry to the source's work. The verb *tend to* implies that SBC's approach is wrong and prospects a counter-assertion on the part of the writers. In fact, the first clause in this example is conceded while the second part is asserted. In essence, the writers negatively present the first approach, including SBC's, and provide a more suitable approach, from their perspective, in the second clause.[3]

The third and fourth stance types are dialogically expansive, *acknowledge* and *distance*. They either show no stance or suggest the writer's withdrawal from the cited information, and in both cases allow alternative interpretations or positions (White 2003). *Acknowledgement* is neutral, no evaluation or comments on authors or their research is present in this type. It represents the default stance in reporting and falls in between positive and negative stances.

Distance tends towards the negative end of the spectrum. It seems reasonable to assume that it has a negative dimension since it shows the writers' withdrawal and/or uncertainty about cited information. If the writer is uncertain about the cited information, then the readers are also likely to be similarly minded and hence it is expansive. Yet, this type differs from *contest*, which clearly shows negativity to cited materials and which is contractive. *Contest* closes the space for the reader's solidarity with *contested* arguments whereas *distance* opens the space for readers' support, scepticism or alternative views.

Acknowledge and *distance* can be realised via various textual resources which include reporting verbs, such as *say, argue, claim* and *assume*, in addition to hedges (Hyland 1994, 2005c). Hedges such as *might* and *perhaps* show "uncertainty to truth of assertion [or attributed information in this case]" and "withhold commitment and open dialogue" (Hyland 2005c: 34-49). The use of these textual resources signals dialogic expansion, as long as they are not accompanied by other markers of positive or negative stance (e.g. *convincingly* argue).

Moving on to citation functions, function analysis is relevant to the role of citations and operates at a higher level than stance. It is as important as stance analysis and, when combined with stance, can reveal how the cited sources act, and are acted upon, in the text. Figure 3.5 below summarises the functional relation.

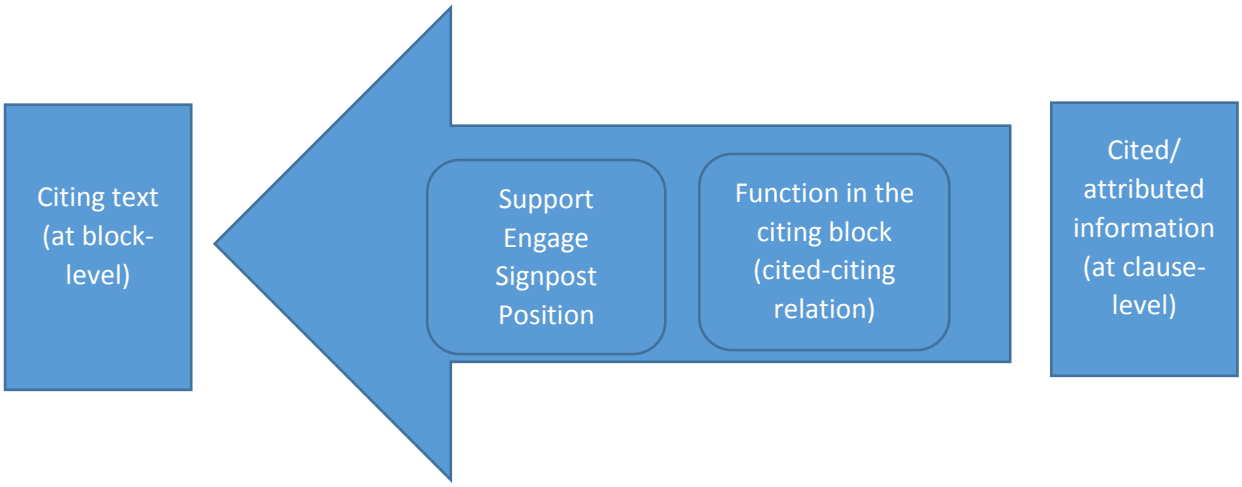

Fig 3.5: Citation functions, the cited-citing relation

Unlike stance, function analysis uses Harwood's (2009) terminology to refer to citation functions. It is important to note that I do not adopt Harwood's methodology for eliciting citation functions for a number of reasons that are enumerated in Chapter 2. As a reminder, Harwood's model is based entirely on interviewing the citing authors. His eleven functions are indeed dictated by the authors, not based on textual analysis. Another reason is that most of the citation classification schemes, including those of Harwood (2009) and Petrić (2007), seem to produce superfluous categories. Therefore, the present study uses only four functions that cover the basic uses and roles of citations and, at the same time, correspond to stance categories. These functions are Support, Engage, Signpost and Position. Their corresponding stance categories are *endorse, contest, acknowledge* and *distance*, respectively.

The focus here is on the observed (as opposed to intended) functions that can be identified at the block-level. Regardless of how (in)accurate the identification is, the ultimate focus is on what the text has to offer instead of what the citer believes or recalls. Harwood investigates authors' motivations for (non-)citing. In other words, each author has a choice between citing or not-citing and Harwood looks into the reasons behind these choices. In the current study, however, it is taken for granted that there is a citation and the choices and methods for coding are different. For example, in the function

"Support", Harwood (2009: 503) elicits this function from his informants based on their perceived reasons for citing particular sources: to support or justify the topic of their research. In the current study, while the same label is used, it is identified in a different way. Analysis in such examples investigates how cited propositions are integrated in Support of the ongoing topic, argument or finding in the citing block, using a close reading method. Example 3.6 briefly demonstrates how Support as a function is identified in this model (the citing statement is italicised).

Example 3.6:

> [1] *During the years 1981-90 the annual rate of deforestation has increased to 16.9 million hectares.*[fn] [2] Unless **this accelerated rate** can be substantially reduced in the near future, the main parts of the tropical forests are bound to disappear by the middle of the next century. [1994_ Döös]

S1 is attributed while S2 is averred by the writer. In short, it could be argued that the citation in S1 is used as background evidence for the averred information in S2. The writers make an anaphoric reference to the accelerated rate of deforestation and thus in a way accept this information and use it as the basis for their argument in S2. So the function of the citation in S1 is to Support the writer's argument in S2. This example is taken from a larger block (see Example 4.38) and included here in order to give a brief account of how Support as a function is identified.

The second function, Engage, is different from *contest* as noted above. Engage is identified at the block-level and relates to the role of the cited information in the citing text. In order to show how this function is identified, consider Example 3.2 in its broader context, block, represented here as Example 3.7.

Example 3.7:

> *[1] First estimates for global resource use* ***were presented by*** *Schandl and Eisenmenger (2006), REF2, REF3 and REF4.* [2] These studies vary in coverage …. [3] They **share a lack of detail** for the Asia-Pacific region, …., and **therefore cannot demonstrate the extent to which** growth in this region has influenced global trends. [2010_Schandl and West]

The writers *acknowledge* the sources of the estimates in S1. Yet, the writers Engage with the sources in S3 and show that the cited estimates *lack detail* for the investigated

context. The writers continue their Engagement and describe the estimate as unreflective of the extent to which growth in the Asia-Pacific region has influenced global trends. In other words, the writers criticise the cited estimates for not considering the Asia-Pacific region. The writers do not imply that the estimates are wrong but instead criticise them for neglecting this important region, in their sense. This is an example of partial criticism. In other words, it differs from other occasions where writers completely reject others' views, findings or arguments. For example, if a writer describes an author's argument as unwarranted, then they clearly indicate that it is unaccepted. The above example shows partial Engagement with sources.

Signposting in citations is the most straightforward function. It plays no role in the build-up of the writer's arguments. It functions to direct readers to external sources and is often signalled by imperatives such as *see, refer to* and *compare*. The identification of this function in the present thesis extends beyond the citing sentence level. For a citation to be labelled as merely Signposting, there has to be no other function in the block. In other words, if a writer uses the imperative *see* to refer to a source in the citing statement but the immediate surrounding context shows the writer's Engagement with the same source, then this citation is taken to show *acknowledgement* at the clause-level and Engagement at the block-level.

The last function is Position. It shows that an argument is cited to show a source's Position, whether it is a prediction, interpretation or viewpoint on the topic under discussion in the block. Consider the following example:

Example 3.8:

> *[1] Although Philippine deforestation rates are relatively low,[1] Lasco and Pulhin (2000)* ***predict*** *that by 2015 carbon storage potential will decrease by 8% due to deforestation, including with protected areas. [2] Across broad parts of the Philippines, the Terrestrial Carbon Group* ***predicts*** *(>75% probability) that at least 10% of additional land area will be converted to agriculture by 2050 (2009).* [2010_Phelps et al.]

In Example 3.8, the writers clearly show that the reported estimates are derived from the cited sources. There is nothing in the citing block to indicate that the cited information in S1 and S2 Support the ongoing argument. The writers are not involved in critical dialogue with those estimates. Therefore, those predictions represent the

sources' Positions. For more examples of the citation functions, see Table 4.1 in Chapter 4.

A very useful approach in studying citation functions is one that considers prospection (and prediction as in Tadros 1985, 1993) as well as encapsulation (Sinclair 2004; Rego 2001). These anaphoric and endophoric relations can be combined with the cohesive devices which link clauses, statements and even paragraphs such as *however, moreover, for example* and many others. The notions of prospection and encapsulation are used in conjunction with the previously mentioned linguistic markers in order to account for citation relations at the block-level.

The combination of stance and function signals at the clause- and block-levels shows the potential for various combinations of form, stance and functions. In other words, coding citations in terms of form, stance and function types facilitates the analysis of each citation at different levels and more importantly reflects how writers manipulate the use of citations to achieve different purposes at the clause- and block-levels. Understanding this variation has various implications for pedagogy, citation theory and future research (see Chapter 7). In summary, the analysis should, and in a pilot study actually did, reveal great variation in the interpretation of citations from the above perspectives.

3.5.2 The block: a new unit for analysing citation functions

The previous literature tends to confine the identification of functions to the citing statements (e.g. Petrić 2007; Snodgrass 2011; Petrić and Harwood 2013; Dontcheva-Navratilova 2016). However, when identifying the function of a citation it is usually necessary to look beyond the immediate citing statement and occasionally the paragraph. Function analysis in this thesis demonstrates the role of citation blocks in determining citation function(s). Consider Example 3.7 where the citation is *acknowledged* at the clause-level (*were presented by*) but involved in critical dialogue in the succeeding two statements. Such examples give an indication of the importance of the notion of the block in deciding on citation role and/or function in the ongoing writer's dialogue.

The obvious candidate for a citation block is the paragraph. However, the current study (e.g. Chapter 6) shows that blocks can, and often do, extend beyond the paragraph. Therefore, in the early stages of function analysis, it was necessary to devise a definition

of the block boundary, at least for the sake of the current thesis. The block of citation is defined here as the immediate context surrounding the citing statement which shows the textual observed function(s) of that citation. In principle, a block could be of any length, even a whole article. In general, however, the shorter the block is, the better it will be for the identification and analysis of citation function(s). As a rule of thumb, the block should end on and include the first sign of citation function. In cases of multifunctional citations, the block can be extended to include the additional co-occurring function(s). However, it was decided that the block should not exceed a maximum of two paragraphs before and two paragraphs after the citing statement. In the majority of cases, the block extends to include the first statement of the following paragraph. I found that even in some papers with extended blocks (e.g. Beckerman and Pasek 1995, see Chapter 6), the blocks do not usually extend beyond two paragraphs before or after the citing statement. Also, for the sake of coherence, it was decided that extending the block any further could cause "incoherent digression" on the exemplified arguments and identification of citation functions (see Hyland 2000: 40, for a similar argument on citation context).

It remains to note that a block can contain a number of citations. For example, some examples that I pick up in Chapter 6 consist of a number of paragraphs. In those examples, there are a number of citations in the block. All the citations are analysed in the block in which they appear. The analysis in Chapter 6 also shows that the blocks are not totally discrete from one another (see Example 6.11 and 6.12 in Chapter 6).

In my opinion, the division of citations into blocks facilitates better capturing of citation functions. Language is highly contextual, and thus to understand how language operates within text, it should be examined in certain contexts. The same concept applies to citation functions: they should be analysed in their contexts of operation. These contextual units are what I term the citation-blocks. The analysis of the conventional and unconventional combinations in Chapters 4 and 6 provide more concrete evidence of the importance of the notion of the block in determining citation functions.

Notes

1- As a student researcher in the same institution in which the project was carried out, I had access to the resources of the IDRD project during the early stages of analysis in 2015).
2- Although this thesis concludes with the idea that the proposed model can be applied in other disciplines and genres (see Chapter 7).
3- Because such examples are uncommon, clause, sentence and statement are used interchangeably to refer to the identified form and stance at the citing-proposition level.

Chapter Four

Data Analysis

4.1 Introduction

Following the definition of the model categories in Chapter 3, this chapter presents and discusses examples of the different analytical categories pertinent and applicable to this piece of research. This is essentially for the purposes of demonstrating how citations are identified and coded. The main argument in this chapter is that citation analysis needs to consider the various meanings conveyed at the clause- and block-levels. In order to understand the bidirectional nature of citing, one has to understand and appreciate how writers present or evaluate citations to achieve different textual purposes.

The current chapter is divided into three sections. The first part of the chapter, Section 4.2, revisits the individual categories of the model and exemplifies each category in isolation from the others. It also provides an overview of how implicit stance has been identified in the current study. In other words, it exemplifies the categories of form, stance and function of citation using explicit linguistic markers available either at the clause-level (in form and stance analysis) or at the block-level (in function analysis). The remainder of this section demonstrates how implicit stance has been identified in the current thesis.

The following section (4.3) is further divided into two sub-sections (4.3.1 and 4.3.2). In this section, the analysis combines the categories of form, stance and function of citation to produce conventional and unconventional patterns of citing. The first sub-section (4.3.1) examines the conventional combinations at the block-level while the second sub-section (4.3.2) shows how the unconventional patterns have been identified: that is when there is a certain degree of mismatch between the identified stance and function categories in a given citation, (e.g. *distance* and Support).

The last section (4.4) revisits some examples, mainly from the unconventional combinations in 4.3.2, to demonstrate how writers draw links between two or more external sources. This section also analyses citations at the block-level, yet the emphasis is not on the role of citations in relation to the writer's argument; instead, it is to show

what role some citations play in relation to other citations from the perspective of the citing writer or cited authors.

As noted in Chapter 3, the current study adopts a theory-based, as well as a data-driven approach in shaping its model. The categories of form, stance and function models (as they appear in Swales 1990; Martin and White 2005; Coffin 2009; Harwood 2009) were adopted and applied to 1186 citations. At the same time, the analysis facilitated revision of the proposed model in order to align the categories. Analysing each of the 1186 citations at the clause- and block-levels to identify the form, stance and function of citation, in addition to the links drawn between external sources, reveals a clear correlation between the form, stance and function categories. That is, as the analysis progressed, it was observed that the clause- and block-level analyses facilitate understanding of how citations are evaluated and also employed in text. Section 4.3 demonstrates the importance of the block in such analysis.

The purpose of this chapter is therefore to exemplify and describe each category. Another aim is to illustrate how the various approaches can be combined in order to understand how citations are used and for what purposes. Further, in Section 4.3, some of the possible combinations are exemplified starting with the conventional (in Sub-section 4.3.1) to the more unconventional combinations (in 4.3.2). The examples in 4.3.1 show correspondence between stance and function categories (see Table 4.2). In other words, the conventional citations are defined here as the ones which act, and are acted upon, in the same way (e.g. *endorse* and Support); taking into account the fact that stance and function operate in reversal directions (see Figures 3.3 and 3.5 in Chapter 3). In Sub-section 4.3.2, the importance of the citation block becomes more evident in the unconventional combinations whereby stance and function analyses yield different results for the same citation.

The unconventional combinations emphasise the need to combine stance and function categories. The reason for this is that if all citations were conventional, as some people might assume, then the practice of aligning stance and function categories will be a mechanical exercise. In other words, we learn more from analysing citations at clause- and block-levels than we do if we analyse citations either at clause-level or even at block-level alone. From this perspective, the adoption of the block as a new analytical

unit to analyse citation function complements the clause-level analysis of form and stance of citation.

In this thesis, it is argued that all the combinations can be observed in academic writing. While some combinations may be less common than others, I would argue that they are all achievable because, although they are connected, they operate at different levels and in different directions (i.e. inward and outward of the citing text). In fact, I would also argue that such combinations (particularly the unconventional) seem to reflect the complexity and creativity of ways in which academic writers use citations to achieve certain purposes.

4.2 The categories revisited: explicit and implicit stance

The proposed model is deconstructed into its constituent levels of analysis in order to provide examples of the individual categories and to demonstrate how each category is identified and coded. Then, Section 4.3 provides examples which show how the categories combine to form one unified model that takes into account different aspects of citation form, stance and function.

To start with, there are three facets of analysis in the proposed framework (as shown in Figure 3.2 in the previous chapter). These facets contain the analytical categories of citation form, stance and function. The analytical categories are exemplified, with some identifying features, in the following table (Table 4.1). It is notable at this stage that the table includes explicit signals of each category, mainly the stance categories. The identification and coding of implicit stance follows in the second half of this section.

Table 4.1: Form, stance and function categories: some identifying features and examples

The category	Some identifying features	Examples
Form categories		
Integral citations	Participant/Subject + verbal or mental reporting verbs e.g. *said, argued, believe*, … etc. Also, other structures such as *According to* or *as found by*, ...	Example 4.1: **Campbell** (1997) **argues** that discourses of vulnerability downplay the resilience of …. [2009_Mortreux and Barnett]
Non-integral citations	Cited statement + reference in parentheses or in footnote	Example 4.2: Deforestation and forest degradation account for 12-20% of global anthropogenic carbon emissions, … **(REF)**. [2010_Phelps et al.]
Stance categories		

Acknowledgement	Some reporting verbs or structures in integral forms e.g. *argue, say, is referenced in,* … and/or unmarked propositions for positive or negative evaluation in non-integral forms, i.e. if the source of the information is presented without evaluation of the source or its information	Example 4.3: First estimates for global resource use **were presented by** REF1, REF2 and REF3 … [2010_Schandl and West**]**
Distance	Some verbs like *claim, speculate,* or structures like *According to,* which are usually accompanied by other signals like quotation marks	Example 4.4: … Rubin and his colleagues **claim** that **'**a major failure of NAPAP is that …**'**. [1995_Herrick and Jamieson]
Endorsement	Verbs or reporting structures like *point out, found, as found in* … adjectives like *compelling, successful, pioneer* or adverbs like *successfully, remarkably*	Example 4.5: The increasing US environmental ethic of the 1960s was explained … in a **seminal** article by Heberlein [1991_Kempton] Example 4.6: **Indeed**, a growing collection of empirical **evidence** now **documents** a range of climate-change-attributed changes in biological processes, including phenology (REF.1), net primary production (REF.2), and inter-specific interactions (REF.3).[2010_Hagerman et al.]
Contest	Some verbs like *overlooks, simplify*, or adverbs like *unconvincingly*, or adjectives like *simplistic*, nouns such as *failure* or negatively comparative structures e.g. *even less case*	Example 4.7: In this regard, SBC's article **is not served by** the **ambiguity** in her use of the word 'alliance' (sometimes meaning an overlapping common interest, at other times a stronger axis of coordination). [1995_Shackley and Skodvin]
Function categories		
Signpost	Imperatives e.g. *see, compare, consider*	Example 4.8: **See** McCright and Dunlap's (2000) analysis and description of the counter-claims made by the conservative movement … [2008_Lahsen]
Position	Hedges e.g. *seem to believe*	Example 4.9: Do Americans feel …? Moral foundations … *By contrast, **it has long been argued** that …*[fn] Within such a moral framework, … *Several current religious and moral writings **take a contrary***

		***position**, arguing that …*[fn] *Some scholars also make the descriptive **claim** that …*[fn] [1991_Kempton]
Support	Agreement with citing text that can be signalled by additive cohesive devices e.g. *Moreover, for example, as found by*	Example 4.10: Nearly 10 years later, analyses of climate science controversy still tend to ignore its deeper socio-cultural roots and …, **as also recently noted by** …[2008_Lahsen]
Engage	Adversative cohesive devices e.g. *However, Nevertheless, yet, on the other hand* … etc.	Example 4.11: *The majority of work discussing the … has been by scientists and journalists without grounding in social science literature (see among many others, REF*). This popular literature tends to highlight financial benefits enjoyed by … **However**, **this paper belies the simplicity** of explanations of the trio's involvement with this movement in terms of calculations of financial gain. [2008_Lahsen]

The categories have been defined in Tables 3.3, 3.4 and 3.5 (see Chapter 3). Table 4.1 outlines some explicit signals of each category and provides brief examples. The first two categories relate to citation form. Integral citations can easily be identified as they usually follow the structure 'Participant/Subject + reporting verb' in Theme position. The reporting verb can be either mental (e.g. *believe, argue, perceive*), verbal (e.g. *said, declared, reported*) or 'active' (i.e. verbs of doing, those which report research not arguments; see Bazerman 1984). Example 4.1 in the table exemplifies reports of indirect speech via use of the verb *argues*. Integral citations can also appear in rheme position and are usually realised by the presence of other reporting structures such as *as proposed by* or *also noted in*.

The second form is the non-integral form of citation. This form follows the structure 'statement + (cited authors in brackets) or statement + superscript number usually referring to a footnote containing full reference entry details'. In this form, the cited author plays no grammatical role in the construction of the cited proposition. In other words, the author's contribution is backgrounded (as argued by Thompson 2001, who found a preference amongst some Agricultural Botany PhD theses for the use of non-integral forms). In Example 4.2, the cited information from the source (indicated by

REF in brackets) is re-proposed by the writers, whilst being attributed to the source. In this way, the sources do not form part of the cited statement, but are attached to it. The different stance and function categories can be realised in both integral and non-integral forms.

The four categories appearing under stance in Table 4.1 are related to writers' evaluation of cited authors or attributed propositions, thus they represent the citing-cited relation (i.e. how the citing text evaluates the cited authors or their argument/research). The table includes the different types from the most neutral (*acknowledge*) to the most critical (*contest*) stances. Example 4.3 in the table simply lists three references in which estimates for global resource use appear. No positive or negative evaluation of the sources is conveyed via the use of the reporting structure *were presented by*.

In the next example (4.4), the writer clearly *distances* themselves from the claim by Rubin and his colleagues. Verbs such as *claim, assume* and *speculate* are signs of *distancing* at the clause-level. *Distancing* may also be signalled by the presence of other reporting structures such as *According to* which shows that the reported (or 'projected', to use Halliday's 2004 term) proposition is from the perspective of the cited source(s) alone.[1] In Example 4.4, the cited proposition is presented as a claim, thus the writer withdraws from sharing responsibility for the sources' argument, which is also quoted. The use of quotations has been extensively studied in the literature on faithfulness in reporting (e.g. Short et al. 2002; Semino and Short 2004). The conclusion from that literature is that quotations are used in a number of ways. For example, writers often quote the words of other researchers and integrate them into their ongoing arguments to *endorse* them (see Example 4.18 in this section). On the contrary, research has shown that writers might "faithfully quote a statement they strongly disapprove of in order to cast maximum blame on the original" source (Short et al. 2002: 352). Example 4.7 includes a quoted word which is used to transfer responsibility and at the same time to cast doubt on the source's use of the noun "alliance" (see Table 4.1).

Both *acknowledgement* and *distancing* are dialogically expansive in that they leave the space open for alternative views by the reader (Martin and White 2005). In *distancing*, the proposition is grounded in "an individualised, contingent subjectivity, that of some external source" (Martin and White 2005: 114). *Distancing* goes "somewhat further

than" *acknowledgement* in that it maximises the space for "dialogistic" alternative voices (ibid: 114).

Some scholars, like Martin and White (2005), raise the issue relating to the difficulty of distinguishing *acknowledgement* from *distancing* in some citations. I argue that while it is difficult to do so, it seems reasonable to make a distinction between the two, whereby the former signals objective reporting by the writer, whereas the latter indicates the writer's subjectivity in their withdrawal and withholding of responsibility for the reported information. To demonstrate the overlap between the two stance types, consider the following example (the linguistic markers of stance are given in bold).

Example 4.12:

> *Still others **argue** that **perhaps** El Salvador is not the environmental wasteland that it has been painted to be (Hecht et al., 2002, 2006).* [2008_Dull]

There are a number of possible interpretations of this example. Expansion in this citation is achieved in different ways. It is clear that Hecht et al. (the source) propose the proposition *El Salvador is not the environmental wasteland …*. One possibility is that Hecht et al. made the proposition expansive by the use of *perhaps*, thus *perhaps* may be assumed to be attributed to the source. However, after consulting the sources, it became clear that they emphasise with certainty that El Salvador is not an environmental wasteland. This leads to the other possibility that *perhaps* is averred by the citing writer. Although this is not the most sensible interpretation since *perhaps* appears in the reported clause, it seems that the citing author used *perhaps* to make the citation more expansive. The first sign of expansion is the verb *argue*. The verb *argue* by itself indicates *acknowledgement*, which is also expansive. However, there are other accompanying linguistic signals which *distance* those making the argument (Hecht et al.), such as *others*. In other words, the writer chose to socially *distance* the arguers thus accepting no responsibility for their argument. This withdrawal is re-emphasised by *perhaps* which suggests uncertainty about the cited information. Therefore, I would argue that this citation exemplifies *distancing* at the clause-level, although it uses a neutral reporting verb.

Drawing the line between *acknowledging* and *distancing* is a non-trivial task as it will always remain a matter of interpretation. This is because citation is highly context-

dependent (Hellqvist 2009) and labelling instances as either *acknowledging* or *distancing* depends on the analyst's interpretation. However, on the positive side, both stance types are expansive and position the writer as relegating responsibility for the attributed propositions. *Distancing* is subjective in that it shows the writer's choice to withdraw from the cited information, thus declining to take responsibility for the attributed message. While the writers also do not take responsibility for an *acknowledged* proposition, it is possible for the same writers to reclaim responsibility for the same proposition at the block-level (see Example 4.36).

As noted in Chapter 3, *endorsement* may be explicit or implicit. Table 4.1 includes instances in which it is clearly signalled by the citing writer. In general, *endorsement* includes any positive evaluation of the cited authors' or propositions either in integral (e.g. 4.5) or non-integral forms (e.g. 4.6). Explicit *endorsement* uses a variety of textual markers, such as reporting verbs (e.g. *point out*), or nouns (e.g. *evidence, success*), adjectives (e.g. *seminal, pioneer*) and adverbs (e.g. *compellingly, remarkably*). Other signals include markers of the sub-system Proclaim (Martin and White 2005) or boosters (as in Hyland 2005c) such as *Indeed, of course, it is well-known that* and many others.

In Example 4.5, the source's (Heberlein's) article is described as *seminal*, thus clearly *endorsed*. The second example in the table (Example 4.6) shows that *endorsement*, just like any other stance category, can be explicitly signalled in non-integral forms of citation, although many people would believe that non-integral citations simply *acknowledge* other sources.

Example 4.6:

> ***Indeed**, a growing collection of empirical **evidence** now **documents** a range of climate-change-attributed changes in biological processes, **including** phenology (REF.1), net primary production (REF.2), and inter-specific interactions (REF.3).* [2010_Hagerman et al.]

The presence of nouns such as *evidence*, accompanied by the verb *documents*, upgrades the status of those sources from simply reporting on various changes in biological processes to being evidential. If a source is described as carrying evidence of any phenomenon, it then becomes contractive and difficult to deny or question by the recipients. It is difficult to challenge a documented piece of evidence, especially when the writers present it as such, therefore these sources are *endorsed*. The contractive

sense in this example is also achieved via the use of resources from the sub-system Proclaim: Pronounce as in Martin and White's (2005) model, such as *Indeed*.

The final stance type is *contest* and it is inclusive of any negative evaluation of the cited source or materials. It is usually signalled by the presence of some reporting verbs or nouns which convey the writer's negative evaluation in integral forms. Consider Example 4.7 in Table 4.1 above (*SBS's article is not served by the ambiguity in her use …*). This example is clearly marked for negative stance (*contest*) at the clause-level. The source's (SBC) article is criticised for the *ambiguity* in the use of the word 'alliance'. The writers argue that the ambiguous use of the term did not serve the cited article. Thus, this example exemplifies negative evaluation of the cited article.

The final four categories in Table 4.1 represent citation functions. Each function is aligned with a stance type. Yet, stance and function categories operate in opposite directions. While stance represents the citing-cited relation in that it shows the citer's evaluation of the cited materials, function is relevant to the role of the cited proposition in relation to the ongoing argument in the block. Therefore, it represents the cited-citing relation. That is not to imply, however, that each stance type merely co-occurs with its corresponding function, for example *acknowledge* with Signpost. It simply means that the functions have been aligned with stance types in order to make the two levels of analysis comparable, so that each citation can be coded in terms of form, stance and function in order to investigate the resulting patterns. Indeed, data has revealed various complexities in citations, whereby some citations have been evaluated in certain ways at the clause-level to achieve unexpected functions at the block-level (see Sub-section 4.3.2 for examples). All the functions in Table 4.1 are analysed beyond the clause-level of the citing statement and presented in their blocks.

The first function is Signpost. It refers to citations which function to guide the reader to external research. This function can be achieved by using imperatives such as *see, compare,* and *consider*. The citation in Example 4.8 in Table 4.1 appears in a footnote. Here, the imperative verb *see* guides the reader to an external source which describes the counter-claims made by *the conservative movement*. McCright and Dunlap's analysis and description of that movement is not important for the build-up or understanding of the writer's argument, therefore it was given in a footnote. In this view, Signposting is a text-external function of citation since it refers readers outside

the citing text, guiding them to external research. The text-external functions relate the cited materials to external sources outside the citing text either to Signpost the readers to external sources or to show external Positions. The text-internal functions on the other hand, are Support and Engage as they explicitly link the cited propositions to the citing text.

Position is the second function of citing. This function can be realised in the block by the presence of some resources from within the sub-system of Entertain (Martin and White 2005) and hedging (Hyland 2005c). Example 4.9 in Table 4.1 demonstrates that the three citations represent different Positions with respect to the writer's question. As the title of that paper (Lay perspectives of global climate change) suggests and as the block demonstrates, these attributed propositions echo different Positions on the topic under discussion. In fact, this function is explicitly signalled as a *position* in the second citing statement, which opposes the cited Position in the first citation. It should be noted that even though the first citation is simply *acknowledged* at the clause-level (*it has long been argued*), it provides one cited Position in the block. The block then compares the three cited Positions in the first, second and third citations.

The third function, Support, is a common function of citations in academic writing. Academic writers cite previous research for a variety of purposes, one of which is to Support their ongoing arguments or findings. This function is usually realised by positive evaluation of the cited proposition beyond the clause-level, when the cited proposition is used to justify, confirm or boost the ongoing argument or finding. At the block-level, this function is usually signalled by markers of agreement or continuation of the argument. These include additive discourse markers such as *Moreover*, *for example* and many others. Example 4.10 in Table 4.1 shows the research noted by the source to be in line with that of the writer. In other words, the source's observation is cited in order to Support the writer's criticism of the analyses of climate science controversy. The writer draws on the source's argument (*as recently noted*) in order to provide additional Support for her argument. The writer thus assimilates (Coffin 2009) a previous finding that is consistent with their criticism in the first clause.

The final function is Engage. As with the other categories, this function is defined in Chapter 3. Unlike Support, it shows citations where the cited source/information is not in line with the citing text. It is inclusive of all instances which demonstrate

disagreement between cited propositions and the writer's averred argument(s). Engagement is usually signalled by adversative cohesive devices (Halliday 1994, 2004) such as *However, Nevertheless* or critical evaluation of some, or all, of the cited information at the block-level.

Example 4.11 in Table 4.1 reflects Engagement with the cited proposition. Engagement is achieved by *However* and *this paper belies the simplicity* in the last statement of the block. The way the cited information is evaluated at the block-level not only shows that the cited literature is inconsistent with the citing study's perspective, but also reveals the writer's criticism of the cited literature (*belies the simplicity of the explanations of* ...). The writer Engages with the explanation of *financial benefits*, which is introduced in the second sentence and then reiterated in the last sentence (financial gain).

So far in this section, I have investigated the analytical categories that are used in the analysis of citation form, stance and function, paying particular attention to explicit linguistic markers, mainly markers of stance. However, as noted in the literature (e.g. Hunston 1994, 2011; Du Bois 2007) and also observed in the current thesis, evaluation of the cited material is usually highly implicit. For this reason, I shall now turn to the investigation of implicit stance as perceived in the current thesis.

Identifying implicit stance is a complex and highly subjective process as it depends on the analyst's "reading position" (Martin 2000: 162) and also leaves open the possibility that the regular readership of the text in question may have varied interpretations (Hunston 2018, personal communication). Identifying implicit stance depends on the analyst's or the reader's interpretation, which in turn is guided by the writer. However, as noted above, it remains a subjective matter. In this thesis, I present how the implicit stance has been identified when the cited proposition is unmarked for positive or negative evaluation of the source or the cited proposition. The following are examples of instances of implicit *endorsement* and *contest* from the data.

Under implicit *endorsement*, I include formulations which do not include explicit markers of positive evaluation (e.g. *point out, demonstrate, seminal, successfully, as noted by*), but those which are recognised by the ways in which cited information is either presented, adopted or 'assimilated' (to use Coffin's 2009 term) into the text. Implicit *endorsement* is usually concealed and identified at the surface of another stance

category such as *acknowledgement*. Generally speaking, implicit *endorsement* is achieved when cited information is presented as true, factual or trustworthy (for example, from a trusted/highly reputable source). Treating these kinds of citations as merely *acknowledging* external research, for example, is an oversimplification of the writer's, as well as the reader's, perspective. Writers draw on various resources to position their readers either to accept or reject cited propositions. The reader (or analyst) then decodes the explicit and implicit meanings being conveyed by the writer. The role of the reader is therefore important not only in the construction of citation (as in Martin and White's 2005 research into intersubjective positioning), but also in the interpretation of the writer's perspectives when citing. This role has been underrated in many citation studies (e.g. in some interview-based function studies which conflate stance with function and the writer and reader perspectives, such as Harwood and Petrić 2012; Petrić and Harwood 2013; see Chapter 2, Sub-section 2.3.4). The following are examples demonstrating how implicit *endorsement* is achieved when presenting cited statements.[2]

Example 4.13:

> [1] *The* ***impact of humans*** *on the global environment* ***is unprecedented*** *(REF1; REF2; REF3;* ***IPCC*** *2007a).* [2] Global climate change, deforestation, dwindling water resources, desertification and loss of biodiversity are **all symptoms of human-accelerated environmental change.** [3] ***These impacts stem from human population growth, and the rapidly expanding level of per-capita consumption*** *(REF).* [2009_McAlpine et al.]

Example 4.14:

> …for two reasons. *First,* CO_2 ***is currently only*** *0.034% ….*[fn] [1991_Kempton]

Example 4.15:

> *Ruminant meat consumption* ***has become*** *a status symbol of the growing affluence of the new consumer societies (FAO, 2007).* [2009_McAlpine et al.]

Example 4.16:

> *World meat consumption* ***increased from*** *47 million tonnes in 1950 to 260 million tonnes in 2005, … (REF).* [2009_McAlpine et al.]

In Example 4.13, the attributed propositions in S1 and S3 are presented as true. At the surface of S1, for example, the writers *acknowledge* the references and show from

where these baseline facts emerge. No explicit evaluation of the sources' argument is given in the clause (e.g. *as demonstrated by*). However, while the writers explicitly *acknowledge* the sources, they imply that their information is true (as something which has started in the past and which continues to have effect). The present simple is often used to report factual information (Shaw 1992) and previous work or arguments that are still valid and *endorsed* (Pecorari 2013; see Section 2.3.2 in Chapter 2). Another indicator of implicit *endorsement* is the fact that the references in S1 include a report from a highly reputable entity in climate change research: that is, the IPCC (the Intergovernmental Panel on Climate Change).

Stance analysis is conducted at the clause-level of citation. Yet, in this example, the block is introduced to show the writer's agreement with the cited information in S1. The highlighted phrases in S2 and S3 (e.g. *symptoms of human accelerated …* and *These impacts stem from human population growth … per-capita consumption*, respectively) all refer to the human impact in S1. The cited propositions in S1 and S3 are explicitly *acknowledged* and implicitly *endorsed*. The implicit *endorsement* is confirmed at the block-level, which is given in full for Example 4.13, since the cited information in S1 and S3 are used in Support of, and are supported by, the averred information in S2. The practice of using citations with multiple stances in this manner in paper introductions is common in the corpus under investigation, particularly in the constellation 1 papers (see Chapter 5).

In Example 4.14, the writers present the cited information as factual. Although there is no explicit mention that the cited information is a 'fact', it is presented non-integrally to give prominence to the information rather than to who found it. Example 4.14 *acknowledges* the source of the cited information in the footnotes, whilst also presenting the information as factual at the clause-level of citation. The use of *is currently only* implies that this information is a given fact and so it is unlikely to be questioned or rejected by the reader. The information is integrated into the text to serve as the first reason for the information in the previous statement, so it Supports and encapsulates its preceding argument (via the use of *First*, see Rego 2001 on markers of encapsulation).

In Examples 4.15 and 4.16, the writers use the present perfect and past tenses to report information which is originally documented in the sources. The citation in Example 4.15 is contingent with the overall argument in the paper that the increasing

consumption of beef is of environmental concern; it is therefore presented as a fact and Supports the overall argument and the succeeding citation in the same paper (Example 4.16). In addition to the fact that this information is presented as factual, it also comes from a trusted source within the discipline, the Food and Agriculture Organisation (FAO).[3] This makes the information more likely to be accepted at least at the clause-level. The writers are re-proposing, and therefore adopting the cited argument in Example 4.15 and the reported fact found in Example 4.16 concurrently.

It is clear from the above examples that such instances act contractively (in Martin and White's terms). In other words, following Hunston's (2000: 187) notions of 'averred assessments' and 'averred facts', I draw a distinction, in attribution and reporting, between *acknowledgement* of factual and true information on the one hand and *acknowledgement* of assessments or opinions on the other. *Acknowledged* facts (e.g. Examples 4.13-4.16) seem to act contractively. This stands in contrast to Martin and White's (2005) identification of *acknowledgement* as always expansive. In other words, it is possible for a writer to *acknowledge* the source of a fact. *Acknowledgement* in this sense is different from *acknowledgement* of opinions or estimates (see Example 4.3 in Table 4.1). When *acknowledging* facts, writers close the space for alternative views, unlike assessments. Put differently, it seems reasonable to argue that the contractive sense of implicit *endorsement* of the cited information overrides the explicit *acknowledgement* of the source of the information. Yet, *endorsement* in such instances is not explicitly signalled. Therefore, it should be identified alongside explicit *acknowledgement*. This is to distinguish citations which use explicit signals of *endorsement* (e.g. Example 4.6 "***Indeed**, a growing collection of ,,, **evidence** now **documents**… "*) from those which are recognised implicitly by the way information is presented, adopted or integrated into the text. Simply put, this thesis argues that *acknowledge* and *distance* are not always expansive (see Examples 4.26 and 4.27 on implicit *contest* in this section). This is because the interpretation of dialogic contraction at the clause-level depends on a number of factors such as the way information is presented, adopted or integrated within the broader block.

One example which shows how the presentation of cited information might take precedence over explicit signalling of the source's arguments can be found in Example 4.17. This example shows implicit *endorsement* via the presence of established sources that are described as *prominent climate scientists* who provide *evidence* of the issues

associated with *the 'desire for more funding' argument*. The *endorsement* of the source is explicit in this example, but it is implicit in terms of their arguments.

Example 4.17:

> [1] Even more challenging to the 'desire for more funding' argument is the **evidence** that some **prominent** climate **scientists** do not believe that lack of money *should* be singled out as the main problem or most important science policy issue. [2] *One of the **most highly respected climate modellers**, Syukuro Manabe from GFDL, **argued** … that it is important to redistribute money to modellers with proven proficiency and to relevant observational work, **rather than simply upping the total global change research budget.**[fn]* [3] *Other **prominent** climate **Scientists** (including REF) **have also argued** that research money needs to be redistributed towards climate modelling ….* [4] **All this points to the fact that** … [1995_Shackley and Skodvin]

The verb *argued* (in S2 and S3) is neutral but the fact that the source in S2 is described as *highly respected* clearly *endorses* that source and potentially their argument as a result. The following citation in S3 provides another similar example since the sources have been introduced as *prominent … scientists*, even though their arguments are explicitly *acknowledged* (*have also argued*). At the block-level, both citations Support, via an act of exemplifying, the first averred statement, and they also lead to the averred *fact*, as presented by the writer, in the last statement (S4). *Endorsing* the sources in this example uplifts their *acknowledged* arguments. While it is true that *endorsement* in S2 and S3 is of the sources, it could equally be argued that it applies to their arguments. The block confirms this conclusion.

Another way of presenting cited information, which contributes to the identification of implicit *endorsement*, is achieved when the cited proposition is integrated into the writer's ongoing argument. There are occasions in which writers use direct quotes (see Example 4.18) or 'assimilate' (to use Coffin's 2009 term) the source's arguments and fuse them into their ongoing text (see Example 4.19).

Example 4.18:

> *Far from lacking information, understanding and sophistication, the public's representation in some respects may be more sophisticated **precisely because it is "**… tied to larger questions of the relations between society and nature" (REF).* [2009_Räthzel and Uzzell]

The writers in this example chose to merge the cited reason with their ongoing argument, thus implying that the quote is the precise cause of the sophistication of the public's representation. In other words, the quotation marks here do not confirm *distancing* as found in other examples, such as Example 4.4 in Table 4.1; rather, the writers chose to report the cause, in the source's own words, and merge it into their ongoing argument. The averred clause which starts with *precisely because it is* demonstrates that the writers believe that the quoted cause is true and consistent with their argument. Another example is Example 4.19.

Example 4.19:

> [1] Developments in health care and social services have greatly reduced mortality and fuelled explosive population growth. [2] *For example, … In most island communities* ***similar*** *developments* ***have put*** *heavy strain on social services and education systems,* ***and have helped to produce*** *a dramatic shortage of skilled personnel in all fields and at all levels.*[fn] [1992_Pernetta]

In the second statement of this example (S2), the writer *acknowledges* the source of the information. This statement includes a self-citation. While this citation explicitly *acknowledges* the source of the information, it implicitly *endorses* the cited message as true. The writer presents the *similar developments* in S2 as something which has led to pressure on the social services and education systems and has *helped to produce a dramatic shortage of skilled personnel …*. Whether attribution occurs in the first, second or the two clauses of S2, the attributed information is presented as true (i.e. development … causing shortage of skilled personnel). At the block-level, this citation Supports the writer's argument in S1. The developments *in health care and social services* in S1 have caused the aforementioned problems. Therefore, function analysis in the block confirms the identification of the implicit *endorsement* in this example.

Implicit *endorsement* is also identified when a writer adopts a methodology, an approach, a concept, a definition or data from another source. In general, writers do not adopt an approach or a definition unless they believe it to be suitable, useful and/or relevant to their work. For this reason, adoption is treated as an indicator of implicit *endorsement* in this thesis. The following two excerpts exemplify implicit *endorsement* via adoption of methodology, methods or somebody's work.

Example 4.20:

> ***An altitudinal correction factor**, based on mean regionalized adiabatic lapse rates for maximum and minimum temperature **as determined by Schulze and Maharaj,[fn] was used** to adjust the mean monthly temperature data ….* [1993_Schulze et al.]

In this example, the writers *acknowledge* the source of the *altitudinal correction factor* and adopt the factor as determined by the source. External work is adopted if believed to be valid. For this reason, it is implicitly *endorsed*.

Another example where external work is adopted is given in Example 4.21 below:

Example 4.21:

> *In developing our analysis, **we have drawn heavily on the work of** N S **Jodha and his colleagues** …[fn]* [1995_Millette et al.]

The same analysis applies to this example. The writers *acknowledge* Jodha's work and also adopt it, thus *endorsing* it implicitly. Other examples of adoption include Example 4.22 (adoption of an approach), 4.23 (adoption of source's figures) and 4.24 (adoption of a definition).

Example 4.22:

> *The supervised **approach** to image classification [fn] **was used** for the three field areas …".* [1995_Millette et al.]

Example 4.23:

> ***The article is concerned with** the old rain forest core region, the Malay Peninsula and the island of Borneo (Figure 1), ….[fn]* [1990_Brookfield and Byron]

In Example 4.23, the source of the figure is given alongside the figure at the top of the page in which the citation occurs. Figure 1 is cited from previous work by Brookfield and colleagues. The writers in this example acknowledge the source and adopt the figure. In other words, the figure is adopted because the writers accept it, even if that acceptance is not explicitly stated at the level of the clause.

The following is an example of the adoption of a term, or in particular, a definition of a term. In this integral citation, the writer clearly shows that he follows Austin's definition. Therefore, the definition is implicitly *endorsed* since it is adopted.

Example 4.24:

> ***Following Austin (2002), I define "antienvironmentalism" as a*** *collection of ideologies and political practices designed to ….* [2008_Lahsen]

The default assumption would be that this citation explicitly *acknowledges* the source of the definition. However, I argue that the definition is also implicitly *endorsed* since it is adopted. Citing the source of the definition (Austin) by itself is neutral and thus shows *acknowledgement*; however, the adoption of the same definition (e.g. signalled by *Following* and *I define*) indicates a role of that definition in the citing text, as opposed to simple mention. Another possible interpretation is that *endorsement* takes precedence in this example. However, I distinguish examples such as 4.24 from others which positively evaluate the source or their work. By way of comparison, consider the following example (4.25), which is a different representation of Example 4.24.

Example 4.25:

> ***Following Austin's compelling definition (2002), I define "antienvironmentalism" as a*** *collection of ideologies and political practices designed to ….*

In this example, I have only added the noun phrase *compelling definition*. However, this makes Example 4.25 considerably different from 4.24. In 4.25, I (or the supposed writer) have included an explicit linguistic marker of *endorsement*: that is *compelling*. Therefore, Austin's definition is explicitly *endorsed* in 4.25. In 4.24, however, the writer explicitly *acknowledges* the source of the definition and, at the same time, implicitly *endorses* that definition since they adopt it.

Moving on, *contest* is usually marked at the clause-level of citations. This by no means suggests that implicit *contest* is unachievable. In fact, there are a number of examples of implicit *contest* in the data. What is noticeable is that implicit *contest* tends to be realised at the surface of an instance of explicit *distance* (as opposed to explicit *acknowledge* which tends to co-occur with implicit *endorse*). Consider the following two examples:

Example 4.26:

[1] *One **may be tempted** to follow Beitz and recognize that, since sovereign nation states are the primary actors in international politics, intercountry redistribution should be viewed as a 'second-best' solution to the global application of the difference principle.*[fn]

[2] There is, **however**, **no reason to believe that this** would lead to any improvement at all in welfare of the worst-off groups in the world. [1995_Beckerman and Pasek]

Example 4.26 is used here to demonstrate how implicit *contest* is identified at the clause-level. To start with, the writers in S1 *acknowledge* Beitz's view of *intercountry redistribution*. It could also be argued that the way Beitz's argument is presented by the writers implies that it should not be followed, thus the writers *distance* themselves from Beitz's argument. The verb *tempted* shows Beitz's argument to be conceded by the writers. S2 acts as a counter-assertion for the attributed proposition/view in S1. Moreover, the determiner *this* in S2 refers to the proposition starting with *intercountry redistribution* in S1. At the proposition-level of S1, no explicit signalling of *contest* is provided by the writers (e.g. *unconvincingly* claims); it seems from first glance that the writers are simply *acknowledging* Beitz's argument. However, in addition to the explicit *acknowledgement* of Beitz's view, the choice of the hedged reporting verb (as in *may be tempted*) implies *contest* of Beitz's argument in S1.

The implied negativity is confirmed by the writers in S2. The use of *however* clearly predicts a contrary position by the writers and signals their involvement in critical dialogue with the source; therefore, it confirms the implicit *contest* that was initially identified at the level of the clause in S1. In short, *contest* in this example is implied, not explicitly expressed. This implication is realised at the surface of explicit *acknowledgement* of Beitz's view and/or explicit distancing (i.e. withdrawal from sharing responsibility for the truth value) of Beitz' view in S1.

To illustrate how implicit *contest* co-occurs with explicit *distancing* at the clause- (not proposition-) level, I draw on one example from Dontcheva-Navratilova (2016: 66). Dontcheva-Navratilova investigates examples such as the following (represented here as Example 4.27) in terms of how writers 'Establish the Niche' move in introduction sections for the presentation of their counter-claims. No analysis of implicit stance is provided by Dontcheva-Navratilova.

Example 4.27:

> [1] *Scollon, for example,* ***not only refutes*** *the sender/receiver model* ***but also claims*** *that it is virtually impossible to define the so-called 'implied reader'.* [2] **From our point of view**, the theory of **'implied reader' is still at least partly valid** because newspapers need to delimit their readership in order to get their share of the market, although they themselves realize that such a reader does not really exist (cf. REF1, REF2)

The citing statement (S1) provides, by means of exemplifying (*for example*), anaphoric reference to previous text. The writers aver that Scollon's views (realised by the verbs *refutes* and *claims*) constitute one example of those who refute the sender/receiver model mentioned [probably] in the preceding statement, which is not included in the excerpt. In this thesis, citing propositions such as S1 are analysed taking into account three aspects.

First, in terms of Example 4.27, this example exhibits the interplay between the writer's evaluation (*writer act*, as in Thompson and Ye 1991) and the source's evaluation (*author act*, ibid). In other words, in this example the writers present the source's view first (*Scollon ... refutes*) and then they introduce their evaluation of the source's argument (via the use of verb *claims* which signals *distancing* of the author's view). In the first clause of S1 (*Scollon ... not only refutes ... model*), the writers present the source's (Scollon's) position: refusal of the sender/receiver model. The source's position is clearly signalled by the writers' choice of the verb *refutes*. Then, in the second clause (also in S1), the same writers who cite Scollon *distance* themselves by presenting Scollon's argument as a claim.

Second, this example provides clear *distancing* of the source's argument at the clause-level. If the citation in the second clause of S1 (*but also claims ...*) is to be analysed at the clause-level alone, then it would simply be an instance of explicit *distancing*. However, since both clauses occur in one citing statement and directly relate to each other, this example implies a negative evaluation of the source's views.[4] The implicit *contest* is achieved in the sense that not only Scollon refutes the model, but also he/she *claims* that it is impossible to define the 'implicit reader'. It seems that the citation in S1 follows a criticism of those who refuse the sender/receiver model and Scollon is cited as one example. The quote is included to cast doubt on the source's use and/or understanding of 'implicit reader' (cf. another source's use of the term 'alliance' in Example 4.7).

The implicit *contest* of Scollon's view is supported by function analysis, taking into account the following averred statement (S2). The citing writers clearly Engage (in critical dialogue) with Scollon's refutation as well as his/her claim. S2 exemplifies emphasised averral as observed in Hunston (2000). The writers made it clear that they aver and hold responsibility for the proposition in S2 (*From our point of view*). In this view, the implied *contest* in S1 is confirmed by the writers' Engagement with Scollon's view in S2.

Compare examples such as 4.27 to Example 4.13 (*The impact …is unprecedented*). In 4.13, the identified stance types are explicit *acknowledgement* and implicit *endorsement* since this proposition is presented as true: something that has happened and has an effect. Regardless of the block, I would argue that this example implies that *acknowledgement* is no longer expansive on such occasions. The same argument applies to the explicit *distancing* in Example 4.27. Both examples carry implicit evaluation of the sources and/or their arguments. Therefore, it could be argued that the implicit stances (*endorsement* and *contest* in Examples 4.13 and 4.27, respectively) override the explicit signalling of *acknowledgement* or *distancing*. On such occasions, the expansive stance types become contractive, as they seem to close the space for alternative views by the readers.

Having said that, one might argue that each citation should be given one stance label. However, it is argued in this thesis that there is occasional overlap between explicit and implicit stances, as exemplified above. This overlap is clearly observed while considering the notions of dialogic expansion and contraction (White 2003) and applying them to cases of implicit stance. In this thesis, it is argued that discussion of the overlap between explicit and implicit stances is necessary since it is possible for an instance of *acknowledge*, for example, to act contractively as in Example 4.13.

The distinction between expansion and contraction is not always as clear-cut as might be expected. For example, when explicit *acknowledgement* co-occurs with implicit *endorsement* (as in Example 4.13), then *acknowledgement* no longer appears to be expansive since '*acknowledged* facts' do not open the space for alternative views, as opposed to *acknowledged* assessments. An alternative interpretation is that implicit *endorsement* takes precedence over the implication of the apparent *acknowledgement*. That is, what seems to be expansive in the clause is in fact contractive. However,

endorsement in such instances is not explicitly expressed at the surface level of the clause (e.g. as in the positive evaluative lexis *point out* and *compellingly*). Therefore, I would argue that examples such as 4.13 (and 4.27) are 'multi-stanced' as they carry explicit and implicit stances. Some readers might perceive Example 4.13 as an instance of explicit *acknowledgement* of the source while others might argue that *endorsement* overrides *acknowledgement* in this example. In this thesis, examples such as 4.13 are treated as instances of *acknowledgement* of sourced facts and so they become contractive since the implied *endorsement* suggests to the reader that they are non-negotiable. The same principle can be applied to the aforementioned example of explicit *distancing* and implicit *contesting* (4.27). *Distancing* no longer becomes expansive since the implied negativity closes the space for solidarity with the source's view. This contractive sense predicts Engagement with Scollon's view, which is achieved in S2.

Overall, this thesis makes a distinction between propositions which use explicit markers of positive stance (e.g. 'Following X's *compelling* definition …' or 'The impact … is unprecedented, *as noted by REF*') and those whereby *endorsement* is not signalled but rather implied (as in Examples 4.13-4.24). The same distinction applies to instances of implicit *contest* as in Examples 4.26 and 4.27. Whenever positive or negative stances are implied rather than explicitly expressed, they are identified alongside the other explicit stance type. By way of exemplification, reconsider the second statement (S2) in Example 4.17 where the writer *endorses* the source but *acknowledges* their argument. Such *endorsement* of the source predicts acceptance of their argument, even though it is presented simply as an argument (as opposed to a *disclosure* or *finding*). The writers' acceptance of the source's argument is articulated beyond the citation (in S1 and S4).

Whenever implicit stance is identified, it is usually confirmed by function analysis. In other words, function analysis complements the identification of implicit stance in the citing statements. For example, the implicit *endorsement* in the second statement of Example 4.17 can be explicitly observed in the block. While the argument of the source (Manabe) is *acknowledged* in S2, it is explicitly *endorsed* (so are the cited arguments in S3) in the block containing S1-S4. Therefore, S2 and S3 Support the writer's argument in S1 (as examples of the *evidence* provided by *prominent scientists*). The implicit *endorsement* of the source's argument in S2 is confirmed by function analysis, which shows the positive evaluation of the prominent scientists, including the cited references in S2 and S3. The following section (4.3) takes the analysis of citation form,

stance and function to another level. The identified categories are aligned to observe the conventional and unconventional combinations of the citation form, stance and function categories.

4.3 The categories combined: conventional and unconventional patterns

In the previous section (4.2), each analytical category was identified, described and briefly explained. The analysis included instances of explicit and implicit stance. From this section onwards, the possible combinations of stance and function will be explored with specific reference to the concomitant analytical levels (the clause and the block). Each citation has a form, reflects a stance, and fulfils a role/function in the citing text.[5] As such, each citation is analysed at the clause- and block-levels to account for how it is evaluated and employed in the block.

The observed combinations will be exemplified starting with the more straightforward instances through to the more unexpected. The analysis also makes reference to some doubtful cases that can be interpreted in different ways. However, such instances are uncommon since the analysis relies on explicit markers of stance and function. In terms of stance, the analysis in the conventional and unconventional patterns counts explicit stance only.

The conventional combinations include citations in which there is an obvious alignment and agreement between the clause-level analysis of stance and the block-level analysis of function. For example, conventional combinations include instances in which the writer *endorses* an argument which in turn Supports their ongoing discussion, finding or position in the block. The proposed model aligns stance and function categories to observe the possible combinations of form, stance and function categories in every citation. Such alignment resulted in the identification of some expected and also unexpected patterns. I have termed the identified patterns ‘conventional’ and ‘unconventional’ combinations.

It is important to reemphasise that this section concerns the specific relationship between the cited propositions and the citing text. It thus addresses the overall role of citations. In order to achieve this, the form and stance of citations are analysed at the clause-level. Then, the same citations are analysed at a broader level (the block) in order to understand their role(s) in the citing text. While both levels are occupied by the writer, the former represents the citing writer’s stance towards the cited materials or

authors whereas the latter represents how the cited information has been employed by the writer: its function in the ongoing discourse.

The main argument in this section is that the introduction of the block as a new analytical unit is useful in the identification of the proposed categories within and beyond the clause-level. We as readers or analysts learn more about citation practices in terms of stance and function at the level of the block than we do from the proposition-level of every citation. The block also facilitates alignment between stance and function categories, even though they operate at different levels and in different directions.

The following sub-sections (4.3.1 and 4.3.2) highlight some of the observed combinations in the current study's data. Then, the following section (4.4) explores a different, yet relevant, relation between the citations themselves, one that is also initiated by the writer but between cited sources: the cited-cited relation.

4.3.1 The conventional combinations

As noted above, form and stance are identified at the clause-level while function analysis is studied at the block-level. All the examples below will be examined at the two levels in order to demonstrate that stance, clause-level analysis is insufficient to understand the role of citations. Although this level of analysis is very useful in understanding the writer's stance towards the source at the immediate sentence-level, it should be conducted alongside a block-level analysis of citation function which can give an indication of the observed role of citations in the citing text.

This multi-level analysis has revealed noteworthy observations with respect to the use of citations. On some occasions, the analyses have demonstrated no great variation between the two levels. This is the case in the conventional combinations. The writers evaluate the source, their work or propositions in ways that are aligned to the function of citation in the ongoing discussion. For example, these conventional combinations occur when the writer *endorses* a source and uses it to Support an ongoing argument, finding or an idea in the block. Alternatively, writers frequently *acknowledge* an existing source in order to Signpost the interested reader to that source, *distance* a cited proposition to show that it only represents a Position and *contest* a source or an argument to Engage (in critical dialogue as in Harwood's 2009 terminology) with it at the block level.

Table 4.2 lists all the resulting combinations following the application of the proposed model. Overall, the proposed model resulted in the identification of sixteen combinations, four of which are conventional and twelve are unconventional. The conventional combinations show a match between the identified stance and function types. The conventional and unconventional combinations can be realised in integral and non-integral forms of citation. The conventional patterns include citations where the citing writers:

- *acknowledge*-to-Signpost
- *distance*-to-Position
- *endorse*-to-Support
- *contest*-to-Engage

Table 4.2: The conventional and unconventional combinations of stance and function categories

The conventional combinations	The unconventional combinations		
endorse-to-Support	*contest*-to-Support	*distance*-to-Support	*acknowledge*-to-Support
contest-to-Engage	*endorse*-to-Engage	*distance*-to-Engage	*acknowledge*-to-Engage
distance-to-Position	*endorse*-to-Position	*contest*-to-Position	*acknowledge*-to-Position
acknowledge-to-Signpost	*endorse*-to-Signpost	*contest*-to-Signpost	*distance*-to-Signpost

In the conventional patterns, the first stance-function combination includes *acknowledge*-to-Signpost citations. These citations include instances where the writers simply *acknowledge* a source to guide (Signpost) the reader to that source. In such instances, the cited source or their argument does not play any prominent role in the writer's ongoing discussion. Citations of this kind perform a text-external function to refer to external research. Example 4.28 below illustrates how this combination is identified in block. In this example, there are two citations, each one followed by a

footnote reference [the citing statements are italicised and linguistic markers of stance and/or function are given in bold in all the excerpts].

Example 4.28:

> [1] Jodha uses a conceptual matrix of five mountain specifities that interact with factors of human intervention …. [2] *Using this framework, Jodha and his colleagues* ***have studied*** *the transformation of mountain agriculture of several villages in the Himalayan Middle Mountains.*[fn] [3] *Shrestha,* ***for example****,* ***used*** *up to 50 indicators …*[fn] [4] The indicators may be visible or concealed and are related to changes in resource base, … [1995_Millette et al.]

Both citations (in S2 and S3) are integral. In the first citation (S2), the writers *acknowledge* other studies which have examined agriculture in several villages in the Himalayan Mountains. The use of the present perfect verb phrase *have studied* indicates neutral reporting of previous research and single studies (as noted by Swales and Feak 2004). Reporting previous studies using action verbs such as *studied, used, examined* carries no stance towards the reported research unless accompanied with other positive or negative stance markers. The *acknowledged* studies appear in a side note to which the superscript number refers (the superscript number is replaced with [fn] in the excerpts).

The second citation (in S3) also *acknowledges* the other source by the use of the non-reporting verb *used.* Both citations Signpost the reader to the cited studies at the block-level. The following table (Table 4.3) summarises the stance and function types identified in the two citations, which share the same block. Both citations exemplify the first combination, *acknowledge*-to-Signpost. The identification of the model categories has only been applied to the statements including citations (S2 and S3). S1 and S4 are averred by the writers.

Table 4.3: Stance and function categories in example 4.28

Sentence number	Stance	Function
S1 – averred	n/a	n/a
S2 – citation (reporting previous research)	*acknowledge*	Signpost
S3 – citation (reporting previous research)	*acknowledge*	Signpost
S4 - averred	n/a	n/a

Another example of agreement between stance and function analyses is Example 4.8 in Table 4.1, represented here as Example 4.29. This example includes a citation which appears in a side note:

Example 4.29:

> **See** McCright and Dunlap's (2000) analysis and description of the counter-claims made by the conservative movement [2008_Lahsen]

In this example, stance and function analyses can be elicited from the clause-level, which also represents the block for this citation. The writer *acknowledges* the source's analysis to Signpost the reader to this external, additional source and this function is achieved via the use of the imperative verb *see*. The surrounding context of that citation does not show any other function, so it only functions to Signpost the reader to the cited analysis and description.

The second group of the conventional combinations include *distanced* citations that serve to showcase the source's Position. In this type, the writer withdraws from the cited information, leaving it up to the reader to decide whether they agree or disagree (White 2003) and crucially the cited proposition is expressed from the perspective of the source alone. Therefore, it is considered expansive as it allows alternative voices at the level of the clause. The following excerpt representing Example 4.30 exemplifies *distancing* of the source's proposition (*Rubin and his colleagues*):

Example 4.30:

> …[1] *Like the ORB, Rubin and his colleagues* ***claim*** *that* ***'****a major failure of NAPAP is that no serious effort was made to define policy-related research priorities and then to shape an appropriate set of projects … to answer these questions in a meaningful way****'***.[fn] [2] For future programmes, …, **the authors suggest** four basic steps: … [3] ***In short****, the goal should be to* ***'****ask the right questions, influence the research agenda, and help ensure that the appropriate priorities are established.****'***[fn] [4] **The authors seem to believe** that … [1995_Herrick and Jamieson]

The first sign of *distancing* in the first citation (in S1) is the reporting verb *claim*. This verb shows that the writers do not share the same opinion and in fact withhold from its reliability. It could also be argued that the writers quoted the cited information in order to allow the reader to draw judgement regarding the cited information and/or the quoted message. The last statement by the writers (S4), *The authors seem to believe that*, also

stresses *distancing* and presents the *distanced* information in S1 as the sources' Position.

The conventional combinations also include instances of *endorsement* (positive evaluation of the source) at the level of the clause and Support at the block-level (the *endorse*-to-Support pattern). At the clause-level, *endorsement* may be fulfilled by the presence of positively evaluative lexis or structures as in the following example which includes explicit markers of *endorsement* at the proposition-level of S2.

Example 4.31:

> [1] A detailed discussion of these explanatory frameworks can be found in Uzzell (2000) but what they have in common is that the kind of theoretical explanation offered by psychology, perhaps with the exception of social representations, is often **ahistorical** and **decontextualized**. [2] ***As Bonaiuto et al. (1996) point out****, … social psychological studies* ***have traditionally treated them*** *as locally centred and limited to a single country.* [2009_Räthzel and Uzzell]

In this example, the first statement (S1) also includes a citation but the focus in this example is on S2 as it shows an *endorsed* citation (*point out*) that acts to Support the ongoing argument in the second part of S1 (i.e. about the explanatory frameworks as being *ahistorical* and *decontextualized*). In the second citing statement (S2), the reporting verb *point out*, combined with the structure *As … point out*, clearly shows the writers' agreement with the argument of the sources. However, the surrounding context is quoted to show the role of the second citation in the citing block. It is inserted in the ongoing argument in a way that Supports the writers' criticism of the psychological explanatory frameworks. The use of citations to Support the ongoing arguments in blocks is a very basic function of citation in general. However, there are other occasions whereby *endorsed* citations act differently in the citing text (see Sub-section 4.3.2).

Endorsement can also be explicit in non-integral forms of citation (as shown in Example 4.5); although some people might argue that non-integral citations simply *acknowledge* other sources. In the following excerpt, the citation (S2) exemplifies explicit *endorsement* of the cited references:

Example 4.32:

> [1] With respect to impacts, the challenge for conservation is that changing temperature and precipitation regimes (REF) **are expected to** interact with other drivers … to influence a range of **biological processes** and ultimately species distribution. [2] ***Indeed**, a growing collection of empirical **evidence** now **documents** a range of climate-change-attributed changes in biological processes, **including** phenology (REF), net primary production (REF), and inter-specific interactions (REF*). [2010_Hagerman et al.]

It is difficult to challenge a documented piece of evidence, as noted in Section 4.2: therefore these sources are *endorsed*. The citation is used in Support of the writers' argument in the second part of S1. The first statement predicts changes in *biological processes* and S2 presents cited evidence which documents that change.

Another example of non-integral citations where the source is *endorsed-to*-Support the writer's argument is given in the following quote (Example 4.33):

Example 4.33:

> ***Further, as has been proposed** (REF), **and demonstrated in this paper**, the perspectives of key stakeholders … **will play** a key role …* [2010_Hagerman et al.]

In this example, both stance and function can be elicited from the clause-level. With respect to stance, the verb *proposed,* in itself, does not indicate *endorsement*. However, the presence of the reporting structure *as has been proposed*, accompanied with *and demonstrated in this paper*, indicates *endorsement* of the source's proposal, which is in turn included to provide additional Support from the literature for the writers' argument. This Support is also signalled by the additive cohesive device *Further* which suggests that this statement follows and Supports the preceding information. The fact that the source proposal is in line with the writers' argument is sufficient to show the function of that citation.

The following example (4.34) is another example of explicit *endorsement* in non-integral forms:

Example 4.34:

> [1] Nearly 10 years later, analyses of climate science controversy still tend to ignore its deeper **socio-cultural roots** and the extent to which it involves a debate

about wider social values, ***as also recently noted by*** *the founding director of the UK Tyndall Centre for Climate Change Research (Hulme, 2007).*

[2] Integrating a strong theory of culture and attending to experiences of modernity, **this article highlights socio-cultural** and political dimensions … on the issue of human-induced climate change. [2008_Lahsen]

Endorsement in this citation is signalled by the reporting structure *as also recently noted by*. In other words, the writer's criticism of analyses of climate science was originally *noted* by the source and then re-proposed by the writer. This re-proposal expresses positivity toward this citation in two ways. First, the writer validates the source's observation. Second, the cited observation is included to Support the writer's observation in the first part of S1, which builds on and confirms the source's note. Moreover, the first statement of the following paragraph (S2) in the same block reiterates the need to consider the *socio-cultural and political dimensions* that have been ignored in the previous research which was conducted 10 years earlier than the citing study.

One might argue that there are two possible interpretations of the source's contribution in this example. The first interpretation is that the citation is reporting as in 'Hulme (2007), the founding director of …, noted that …' which suggests that Hulme was the founding director. The other possibility is that Hulme was not the founding director. This possibility is non-reporting in that the interpretation here perceives (Hulme), the cited source, as the source of information 'that the founding director (who is another person supposedly) had noted the observation'. In other words, the information that was noted by the founding director appears in this source (Hulme 2007). However, with reference to the source, it became clear that the source (Hulme) was actually the founding director of that centre which supports the first interpretation. This gives prominence to the issue of presentation and faithfulness in reporting as raised by Short et al. (2002) and averral and attribution and responsibility (see Sinclair 1987; Hunston 2000). The fact that such instances can be interpreted in more than one way supports Short et al.'s observation on the role of the reporter in the reporting context and the confusion that the nature of reporting [particularly in non-integral forms] might cause for the reader who does not usually have access to, or the time to check, the original source (see Coffin 2009 for a similar argument).

The last group of the conventional combinations consists of *contest*-to-Engage citations. This reflects any negativity by the writers towards the cited author(s) or information. This negativity has to be marked at the clause-level in integral or non-integral citations. The writers' negative evaluation of the source is identified in the clause- and block-levels in this combination. The following is an example of *contest* in citation (4.35):

Example 4.35:

> [1] The Geophysical Fluid Dynamics Laboratory (GFDL), one of the premier climate modelling centres in the USA and world-wide, **actually suffered** a 10% **decline** in its base funding from 1990-93.[fn] [2] **This is difficult to understand from SBC's analytical perspective**, in which climate modelling centres are presented as a key beneficiary of **increased** research funding. [3] As important as the absolute funding *per se*, is the change in the relative funding of different research fields. [4] *This suggests that* ***rather than*** *treat the scientific community's priority setting as sovereign (****as SBC tends to[8]****),* ***an important role for the interpretative social sciences is to*** *explore the debates and arguments within the scientific community over … and how money should be spent.* [1995_Shacley and Skodvin]

In the citation appearing in the fourth statement (S4), the source's (SBC's) approach to treating the *scientific community's priority setting as sovereign* is disparaged and the writers call for an alternative approach which focuses on exploring these *debates within the scientific community* and redistribution of money. The averred clause in the second part of S4, as well as the other statements, demonstrates Engagement with the source's (SBC) approach. The writers' approach is presented as a more *important* and possibly suitable alternative approach than that of the source (SBC). The use of *rather than* and *tends to* in this particular citation predicts a shift from the source's approach to a more important approach which considers the *important role for the interpretative social sciences*, one which explores the arguments and debates within the scientific community as averred by the citing writers. Also, the fact that the *premier* centre *suffered* a *decline* in funding contrasts with SBC's presentation of *climate modelling centres* as key beneficiaries of *increased* funding, according to the citing writers at the block-level.

To summarise, the above instances represent the conventional combinations of stance and function of citation. In such instances, stance and function analyses are consistent with one another. The block-level analysis has so far revealed the importance of citation

function analysis. This analysis becomes more important in the unconventional combinations where each line of analysis produces different results.

The block has also demonstrated the importance of aligning the stance and function categories. In this regard, it should be noted that if all citations are conventional (as some might expect), then there would be no point in comparing stance and function categories at the clause- and block-levels. However, since the combinations are not always conventional, it is essential to align and compare the various categories. Such alignment reflects how citations may be evaluated to achieve a different function in the text. The following sub-section illustrates the importance of the block in identifying the stance and function in the unconventional combinations.

4.3.2 The unconventional combinations

The proposed model, in providing a new way to investigate the role of citation in texts, shows that in some citations there is a mismatch between the apparent stance and function. This mismatch resulted in the identification of twelve unconventional patterns (see Table 4.2). This sub-section discusses examples of various possible combinations, as observed in the data.

The twelve unconventional patterns can be observed in integral and non-integral forms of citation. In fact, one aim of the proposed model is to demonstrate that stance and function are different, yet related concepts, as opposed to various studies in the literature which tend to conflate the two (see Chapter 2, Sub-section 2.3.4). Presenting stance and function as two different levels therefore demonstrates that writers have a variety of choices at each level. The writers' choices sometimes match at the clause- and block-levels producing conventional combinations and sometimes mismatch, producing unconventional combinations. Consider the following example (4.36):

Example 4.36:

> [1] *In the early 1970s, American environmental sociologists* ***predicted*** *that national efforts … **would "run** head-on into many traditional values and time-honored practices" (REF).* [2] **This paper confirms their prediction, revealing** … [2008_Lahsen]

Example 4.35 includes an *acknowledged*-to-Support citation. In the first statement (S1), the attributed information is shown as a prediction. Prediction exemplifies

acknowledgement (whereas speculation is an indicator of *distancing*). Prediction and speculation are both expansive. Therefore, in this example, the writer introduces the cited proposition expansively to the reader who is given a choice of whether they agree or disagree with the cited prediction. At the block-level, however, the case is different. The writer supports and confirms the source's prediction. Also, by confirming the source's prediction, the writer closes up the space for alternative views, thus making it contractive (as definitely true). One way or another, the cited prediction thus Supports the following statement (S2) as it is not only confirmed by the writer, but provides a background basis for the writer's argument in the second statement and therefore Supports it, even if not explicitly stated. In other words, the source's prediction in S1 is in line with that of the writer in S2 and even though it is presented before the writer's view in S2, it is included in such a manner to provide evidence for the writer's finding. In other words, the writer's finding in S2 had already been predicted by the source in S1. Therefore, not only S2 *endorses* and supports S1, the cited prediction in S1 is cited to Support the writer's finding in S2.

The question that needs to be asked, as a result, is whether Support, as a citation function, is contractive in this case. I believe it is. In other words, the writer presents the cited information expansively in S1 (*predicted*) and contractively in S2 (*This paper confirms their prediction*). This question arises with any combination of course (e.g. in *distance*-to-Engage citations), yet there are occasions when it seems inapplicable (for example, in *acknowledge*-to-Signpost combination). The function of Signpost does not seem to lend itself to the dialogic perspective since it plays no role in the build-up of the writer's argument in the block.

While Martin and White's (2005) notion of dialogic engagement has been applied to implicit stance at the clause-level (e.g. *acknowledgement* of facts versus *acknowledgement* of assessments), this thesis argues that the same notion can be applied when considering citation functions at the block-level. However, there are two important aspects that need to be considered when applying dialogic engagement to citation functions. First, there needs to be a consensus on the meanings of expansion and contraction before one attempts to identify them in citation function analysis. Second, and more importantly, the analysis has to take into account the different kinds of audiences that some texts are directed to. For example, the function Engage might be read as expansive or contractive depending on the kind of audience and analyst who

interpret its meaning in a given citation. If a writer Engages, in critical dialogue, with a source's view, then that Engagement can be seen as contractive when it closes the space for supporters of the source's view. At the same time, the same Engagement could be seen as expansive for the lay audience as it seems to invite alternative views. See Example 4.39 where the Writer's Engagement with the cited estimates in the block could be seen as expansive or contractive depending on the audience. One possible interpretation is that the writer's Engagement is contractive for the academic audience who might support the cited estimates. In other words, the writer criticises these estimates in S2 and S3 and therefore closes the space for reader support of the estimates. At the same time, another possible interpretation is that the writer's Engagement is expansive for the lay audience as it bids for alternative estimates or views. In other words, the cited estimates are not suitable, as suggested by the writer, and therefore one might come up with alternative estimates. In order to illustrate how the notions of dialogic expansion and contraction can be analysed beyond the proposition-level of the citing statement, consider the following invented example:

> [1] ***Sinclair (1987) argues*** *that the notions of averral and attribution are important in identifying textual and intertextual responsibility.* [2] ***However***, ***there is evidence in the literature (e.g. Martin and White 2005****: 136) that these two notions are "secondary to the issue of how the authorial voice is positioning itself with respect to the anticipated reactions and responses of the audience which is being construed for the text".*

Both statements (S1 and S2) include citations. Martin and White (2005) apply the notion of dialogic engagement only at the proposition-level. For example, according to Martin and White, S1 is expansive because it simply *acknowledges* an external view which could be right or wrong. S1 exemplifies the Attribute system whereby the responsibility for a proposition is relegated to an external source.

In terms of S2, Martin and White (2005) would treat it as dialogically contractive drawing on resources from the sub-systems Concur (e.g. *However*) and Pronounce (e.g. *there is evidence … that*). In citation analysis, Martin and White would analyse the two citing statements in isolation, individually at the clause-level of each statement. In other words, they do not look into how the second statement (S2) relates to the first (S1) to account for the function of S1 in the discourse. In the current thesis, while both statements include citations, S2 is taken as indicating the function of S1.

If the two citations are analysed in isolation of each other, then S1 would be expansive since the verb *argues* opens the space for alternative views (White 2003). It is also possible to argue that there are two possible interpretations of S1 depending on the kind of audience it is directed at. In other words, Sinclair is a highly respected source in the Birmingham school of applied linguistics. Therefore, if taken out of context this statement could be perceived as contractive, for example by the followers of the Birmingham school, even though it includes resources of expand (e.g. *argues* as opposed to *points out*). At the same time, there is the lay audience who will definitely see it as expansive because it is only *Sinclair argues* as opposed to *Sinclair demonstrates.*

Analysis of citations in the current thesis extends beyond the level of the citing statements. For example, in order to analyse citations like S1, I investigate how the citation is presented in relation to its immediate surrounding context (the block which includes S1 and S2 in this hypothetical example). S2 clearly shows the hypothetical writer to Engage with Sinclair's argument in S1. This Engagement is achieved by *However* and *evidence* in S2. In other words, *However* predicts a contrary argument in S2. Also, the writer describes the cited argument in S2 as including or demonstrating evidence (*there is* ***evidence*** *in the literature … that*). Therefore, the cited argument in S2 is *endorsed* by the writer and is contractive.

In function analysis, S1 exemplifies the unconventional *acknowledged*-to-Engage pattern. The writer presents Sinclair's argument expansively at the clause-level of S1. If we take S2 into account, it becomes clear that the writer Engages, in critical dialogue, with Sinclair's argument in S1. The question that needs to be asked then is whether Engage is contractive or expansive in such examples. As argued above, this depends on how Engagement is defined and the kind of audience to whom the text is produced. In other words, the current thesis makes a distinction between academic and lay audience and/or between supporters and opponents of the source's view. In the aforementioned example, Engagement, as a function of citation, could be seen as contractive for the academic audience or those who follow Sinclair's view about the importance of averral and attribution (including myself). The hypothetical writer Engages with Sinclair's argument, arguing that there is evidence in the literature to show that Sinclair's argument is not necessarily true. Therefore, the readers who support Sinclair's views are positioned to think that Sinclair's argument is not necessarily true. For the lay

audience, who may or may not support Sinclair's view, the function Engage in this block is definitely expansive. The writer positions the lay audience not to accept Sinclair's argument, but instead to think of alternative arguments along the lines of Martin and White.

Below are two more examples showing *acknowledgement* of the sources but in Support of the writer's arguments. In the first example (4.37), the writer *acknowledges* the source of an opinion and/or an argument to Support their ongoing argument.

Example 4.37:

> [1] True descriptions of empirical objects are inexhaustible and one is not better than another for all purposes. [2] ***Waismann*** [fn] ***argued*** *that the 'open texture' of scientific findings makes it impossible to provide complete descriptions of most empirical concepts*. [3] **For this reason**, empirical claims can rarely be verified completely. [4] More tests can always be demanded and additional descriptions can always be given. [1995_Herrick and Jamieson]

At the clause-level, S2 simply *acknowledges* the source of the argument (Waismann). No positive or negative stance is assigned to the source or their argument. At the block-level, however, this argument Supports the citing writers' argument that the scientific descriptions of empirical objects are 'inexhaustible', and the following statement (S3) is presented as a consequence of this argument, *For this reason*. In this view, *For this reason* in S3 means because the open texture makes it impossible, as Waismann argued, the writers in S3 and S4 argue that more tests are likely to be demanded. *For this reason* in S3 assumes the truth of the source's argument in S2. In short, the writers *acknowledge* Waismann's argument in order to Support their own argument at the block-level.

The following is a non-integral example of the *acknowledge*-to-Support combination. Example 4.38 contains two citations (in S2 and S3) which are cited in Support of the writer's argument in the block.

Example 4.38:

> Tropical deforestation. [1] In the tropics deforestation **has increased alarmingly** during the past decade. [2] ***In 1980 it was estimated*** *that the loss of the total (open and closed) tropical forest was about 11 million hectares per year*.[fn] [3] ***During*** *the years* ***1981-90 the annual rate of deforestation has increased*** *to 16.9 million hectares*.[fn] [4] **Unless this accelerated rate** can be substantially reduced in the near future, the main parts of the tropical forests are bound to disappear by the middle of the next century. [1994_Döös]

In the first citation (S2), the writer *acknowledges* the source of the information regarding the estimated total loss of tropical forest, that is, FAO and UNEP (1981). Also, in the second citation (S3), the source (FAO 1991) of the information regarding the annual increase in deforestation is *acknowledged*. However, the information in S3 is presented as factual (i.e. *annual rate … has increased*). The increase in deforestation rates is reiterated in the last averred statement (*Unless* ***this accelerated rate*** *…*) which shows Support between S4 and S3 as well as S2. Therefore, this fact, along with the trusted source of the information (FAO), signals implicit *endorsement* towards the second citation (S3). Both citations in S2 and S3 Support the first and fourth averred statements that *deforestation has alarmingly increased* and *unless this accelerated rate* is reduced (the accelerated rate refers to the increase from *11 million hectares* in S2 to *16.9 million hectares* in S3), the majority of the remaining forest is *bound to disappear*. The first citation (S2) is unconventional as it *acknowledges* the source of the estimate (*it was estimated*). In S4, the phrase *this accelerated rate* assumes the truth of the source's estimate in S2 and also S3's finding. Therefore, S2 exemplifies the *acknowledge*-to-Support pattern.

In sum, both S2 and S3 Support the averred argument in S1 about the alarming increase in deforestation. S1 introduces the alarming increase as presented by the writer. Then, the cited arguments in S2 and S3 encapsulate and elaborate on S1's argument. Finally, S4 provides an anaphoric reference to the arguments in S1, S2 and S3 and also re-evaluates as true the *acknowledged* argument and fact in S2 and S3, respectively. In other words, the writer accepts (supports) the cited statements in S2 and S3 and at the same time uses them in Support of the argument in S1 about the alarming increase.

At the clause-level of S2, the writer simply *acknowledges* the source of the estimate expansively. However, if we expand the block to include the first and last statements in the block, it becomes obvious that S2 and S3 are cited in Support of the writer's

argument in S1 and also in S4. Therefore, Support as a function is contractive as it closes the space for alternative views. Overall, this example shows mutual support between the writer and the cited sources at the block-level.

Expressing the same stance but fulfilling a different function (*acknowledge*-to-Engage) is Example 4.39:

Example 4.39:

> [1] *First estimates for global resource use* ***were presented by*** *REF1, REF2, REF3 and REF4.* [2] **These studies** vary in coverage …. [3] They **share a lack of detail** for the Asia-Pacific region, …, **and therefore cannot demonstrate** the extent to which growth in this region has influenced global trends. [2010_Schandl and West]

The writers in this example *acknowledge* the presenters of the estimates at the clause-level. However, they are involved in critical dialogue with the sources in the block; therefore, they inevitably Engage with them. The sources in S1 are criticised and shown to be lacking in depth and irrelevant to the region being examined in the citing study (the Asia-Pacific region). Also, the same sources are criticised for their inability to demonstrate the growth's impact of that region on global trends. Therefore, this is an example of the *acknowledge*-to-Engage combination.

The following is another unconventional example which shows the same function but with a different stance type (*distance*-to-Engage).

Example 4.40:

> [1] ***It has long been speculated*** *that the high rate of deforestation in El Salvador is a function of population density (REFs).* [2] This **simplistic** and **controversial** explanation for the environmental predicament of El Salvador relies on a classic Malthusian resource scarcity model. [2008_Dull]

The argument in S1 is attributed to the source. This attribution is signalled by the reporting phrase *It has long been speculated that*. The presentation of the sources' view as a speculation suggests *distancing* at the clause-level. However, the description of this view as *simplistic* and more importantly *controversial* in S2 clearly suggests that it is open for criticism and allows for other views. The cited speculation is *distanced* at the clause-level of S1. The writer Engages with the same speculation in S2. Therefore, this citation is *distanced* for the purpose of Engagement. Engagement could be seen as contractive on this occasion if it is considered to be showing the sources' view to be

problematic, thus closing up the space for supporting views of the sources' Position. Engagement could also be seen as expansive. Because the sources' explanation is *controversial*, as the writer describes it, then the space is open for new alternative positions. Describing the source's explanation as *controversial* also suggests that other views co-exist or are at least likely to exist after reading the source's explanation. In short, this citation is expansive at the clause-level and contractive, or expansive, in the block depending on which audience it is directed to (i.e. contractive to supporters of the cited explanation or expansive to lay audience to whom this criticism at the block-level may trigger alternative explanations).

The following example *acknowledges* the sources of the argument in S3 to exemplify one Position, among many others in S5 and S6. This example is included in Table 4.1 as Example 4.9. Here, it is used to reemphasise the difference between stance and function identification and to illustrate another unconventional patter: *acknowledge*-to-Position.

Example 4.9 revisited (see Table 4.1):

> … [3] *By contrast,* ***it has long been argued that*** *...*[fn] …[5] Several current religious and moral writings **take a contrary position**, **arguing that** …[fn] [6] Some scholars also make the descriptive **claim** that …[fn] [1991_Kempton]

Regardless of the links drawn by the writer between the sources in the block (e.g. S5 showing citation-engagement with S3), the writer *acknowledges* one Position in S3. This *acknowledgement* is signalled by the phrase *it has long been argued that*. The key word in this phrase is the verb *argued*, which simply *acknowledges* attributed propositions. The writer is not committed to the attributed positions in S3, S5 and S6. The two positions in S3 and S5 are *acknowledged* (*argued* in S3 and *take a contrary position* in S5) while S6's position is *distanced* (*claim that*). In this view, the cited claim in S6 is conventional (*distance*-to-Position). However, the cited arguments in S3 and S5 are unconventional (both are *acknowledged* to show different Positions). All the citations in this example are expansive. *Acknowledge* and *distance* are two expansive stance types as they leave the space open for alternative views. Along these lines, Position as a citation function is always expansive as it simply states sources' views, to which the writer is uncommitted. By way of comparison, consider the function Engage in the following example (4.41).

In the following example, the writer is committed to the cited argument at the clause-level as he *endorses* the source in S1. However, the same writer is uncommitted (and so are most island societies, as presented by the writer) to the implications or presuppositions of the attributed information in S1 at the block-level (including S2-S5). *Endorsed* citations can be involved in critical dialogue with the citing text as in the following *endorse*-to-Engage example. The Engagement in this example is with the presuppositions of the attributed information.

Example 4.41:

> [1] ***O'Collins points out*** *that emigration places* ***severe*** *strains on migrants, host communities, and on those who remain behind.*[fn] [2] **However**, **most island societies view emigration** as a temporary **solution to transient** problems. [3] It is difficult to predict what will be the long-term attitude of island populations to migration that is undertaken in response to degrading environments and climate change.
>
> [4] **Most of the small island countries** are widely regarded as not being economically viable **without external assistance**. [5] **Their survival is closely tied to remittances provided by non-residents**, aid, and development funds from the international donor community. [1992_Pernetta]

The cited statement is *endorsed* at the clause-level, signalled by the positive reporting verb *points out*. Nevertheless, the following statement by the writer shows Engagement with O'Collins' argument by *most island societies*. Yet, at the same time, the writer avers the contrastive cohesive device *However* which signals a change in the writer's position. In other words, *However* Engages with the 'severe' part of O'Collins' argument. Engagement here is with the implications of O'Collins' view. It is fulfilled in two ways; first, by the writer who avers the *However* (which in turn contradicts the severe part of O'Collins' argument) and, second, by *most island societies* who view emigration as only a temporary solution (not a severe strain) to transient problems. In this view, S2 indicates, from the perspective of *most island societies*, that the strain is not as severe as O'Collins suggests. The *most island societies* view in S2 is picked up by the writer in S4 and S5 to show that many small island countries are actually benefitting economically from the migrants (*non-residents*).

In this respect, it should be noted that Engagement in this thesis does not only refer to disagreement with or rejection of the cited information, it also means presenting a perspective that is completely different or contradictory to the source's perspective,

especially when the writer contributes to the Engagement as in the above example. The writer here does not reject O'Collins' argument in the block; rather, he Engages (along with most island societies) with the implications or presuppositions of the source's view. The citation in S1 is contractive at the clause-level since it is *endorsed*. However, it is expansive in the block since the writer also acknowledges another alternative view, thus indicating that O'Collins' view co-exists with others. Thus, the reader has a choice regarding which view to adopt (the view of *O'Collins* or *most island societies*).

Example 4.42 below exemplifies another unconventional combination: the *contest*-to-Support pattern.

Example 4.42:

> [1] The Marshall Institute Trio **has influenced top-level political decision makers' positions** on the climate issue. [2] ***Their 1989*** *Marshall Institute* ***report*** *[REF]- which, …,* ***was not peer-reviewed - have [sic] found receptive audiences****, especially among conservatives and some representatives of fossil-fuel interests.* [3] **According to the Marshall Institute itself, the trio's 1989 report "dominate[d] the White House thinking" on climate change during George H.W. Bush's presidency** (1988-1992). [4] It is also said to have "provided the foundation for the [Bush] administration's resistance to …. " (REF). [5] When asked to comment on the IPCC's first (1990) assessment report, **President H.W. Bush preferred to defer to the 1989 report** by the George C. Marshall Institute, saying: "My scientists are telling me something very different" (REF) [2008_Lahsen]

The highlighted relative clause in S2 (*which, … was not peer-reviewed*) reflects criticism of the trio's report: *contest* at the clause-level. However, the function of the whole citing statement in S2 is to Support the writer's argument in S1: that the trio's work has *influenced top-level political positions* even though the writer does not approve the trio's ideas in the text and is critical of their report in S2. The phrase *found receptive audiences*… in S2 provides an anaphoric reference to the averred phrase *has influenced top-level political decision makers' positions* in S1. The influence of the report (which is cited in S2) is in line with the writer's arguments in S1, S3, S4 and S5.

Even though the report is *contested* at the clause-level of S2, the overall function is to Support the general idea in the block, that the trio's work has enjoyed high-level support. Put differently, the writer *contests* the content of the cited report and at the same time uses the report as evidence to Support the ongoing contention that it was (unfortunately) believed in several quarters. The first averred statement (S1) prospects

(Sinclair 1991; Rego 2001) how the rest of the block should unfold: to elaborate on and demonstrate how the trio's work has been politically influential. The following citations in this block (S2, S3, S4 and S5) are cited as supporting evidence of the argument in the first statement.

The aforementioned examples emphasise the importance of considering the block in citation analysis. This is due to the fact that it can reveal the extent to which clause-level analysis truly reflects the role of a given citation. Also, such block analysis could be helpful in understanding the citation value, an area of research that continues to be underrated in favour of citation counts. Scientific impact, for instance, has long been measured via citation counts. In order to assess somebody's research contribution, the question that has always been asked relates to the frequency or number of mentions of a given citation, and not to the value of each citation or mention. An *acknowledged* resource which shows that a writer has consulted a wider literature and aims to guide the reader to that source is not as valuable as an *endorsed* source that Supports the main argument in the block. Simple citation counts or automatic retrieval systems fail to distinguish the various roles of citations in context. For example, such systems would treat the cited references in Example 4.39 as being *acknowledged*, whereas the wider context shows them to be lacking in depth. This issue is revisited in Chapter 7.

One of the main issues during the analysis of stance and function of citation relates to the interplay between averral and attribution. One example relating to this issue is Example 4.43 (overleaf) where the citations in S1 to S6 might be interpreted as either reporting or non-reporting. The issue lies in whether the information in the non-integral citation is originated by the source (reporting) or whether the source is intermediary and in turn reporting information from another source (non-reporting).

In terms of reporting and non-reporting, the third citation (S3) might be interpreted in two ways: 1) [the sources] assert that no firm evidence has … (reporting) or 2) No firm evidence has yet emerged [where the cited sources are not the ones making the assertion, thus non-reporting]. This is clearly a matter of presentation, and interpretation as well. Similar tensions have been raised by Swales (1990). The *may be* and *can lead to* may well be originated by the sources or be otherwise introduced by the citing writer. However, this should not affect the analysis as the analysis is based on the identification

of the linguistic signals, even in some doubtful cases where it is unclear which parts are averred and which are attributed.

Example 4.43:

> [1] *It is **well known** that increased levels of carbon dioxide and other gases, such as methane, nitrous oxide, and chlorofluorocarbons in the atmosphere **can lead to** the greenhouse effect.*[fn] [2] ***But claims** that a global warming has already begun **may be premature**.*[fn] [3] ***No firm evidence has yet emerged** that the world has become significantly warmer over the past ten or even hundred years.*[fn] [4] ***Even if such changes will soon be documented**, **it will be difficult to** separate natural climatic cycles from any greenhouse effect and the role of clouds and oceans on global temperature changes is imperfectly understood.*[fn]
>
> [5] In the event that the greenhouse effect takes hold, tropical deforestation will be **only partly to blame**. [6] *Deforestation **accounts for less than 20%** of greenhouse-gas emissions.*[fn] [7] Carbon dioxide from the burning of fossil fuels … is the largest component of greenhouse gases. … [1991_Smith et al.]

The first citation (S1) is *endorsed* (*well-known*), the second (S2) is *contested* (*premature claims*), the third (S3) and fourth (S4) citations explicitly *acknowledge* and implicitly *endorse* the cited propositions (presented as true), respectively, and so is the fifth citation (S6), in which the cited information is presented as a Supporting fact of the averred information in S5. All the citations Support the writers' argument in the block that deforestation should not be the only factor to blame. The way in which citations relate to each other, as opposed to the citing text, is discussed in the following section.

4.4 The cited-cited relation

In the previous two sections (4.2 and 4.3), each example was analysed in terms of two relations: the citing-cited and the cited-citing. The former relation (in Section 4.2) represents the analysis of stance in citation practices and concerns the citing writer's evaluation of the cited sources, their arguments or findings. The latter relation (in Section 4.3) relates to the role of the cited materials in the citing text. In other words, it looks into how the cited materials are employed by the writers to achieve certain functions which are either internal or external to the citing text.

In this section, some examples from the previous two sections are revisited in order to draw attention to a different, yet related, relation: the cited-cited relation. Before

introducing this relation, it is useful to recall the visual representation of the different relations, as presented in Figure 3.1 (represented here as Figure 4.1).

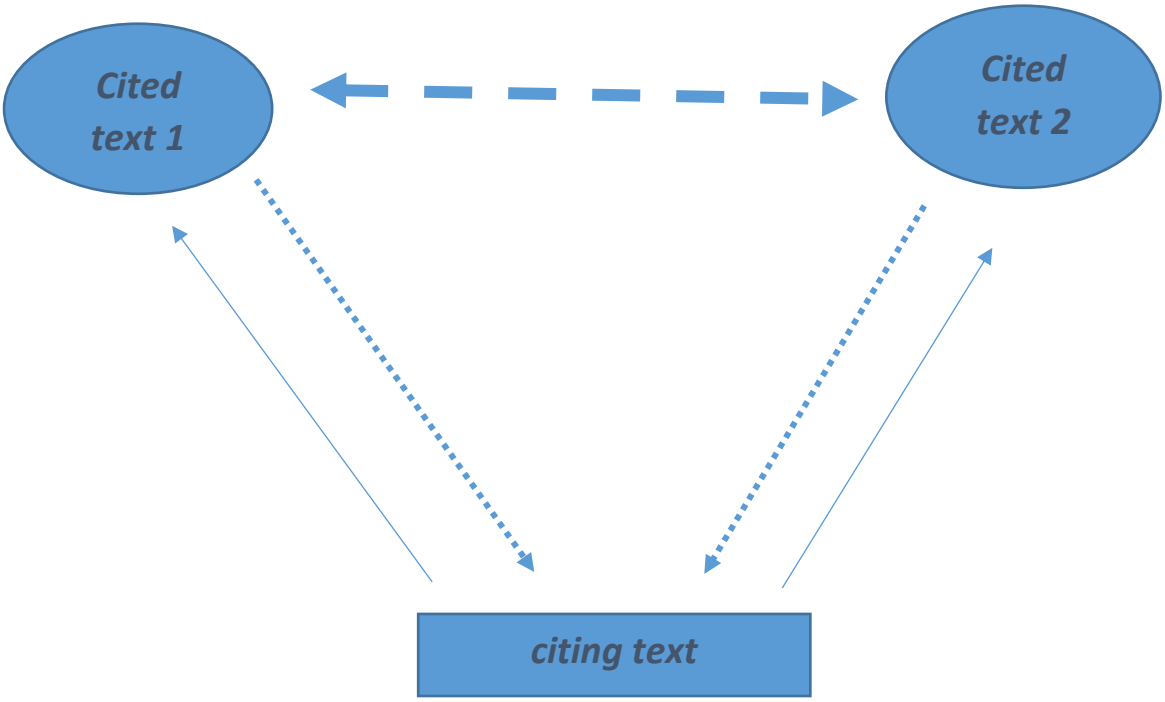

Figure 4.1: The relationships between the citing and cited texts

Figure 4.1 contains the three relations identified in the current thesis between citing and cited texts. In summation, the solid lines represent the citing-cited relation (stance) whereas the cited-citing relation (function) is represented by the dotted lines. The dashed line represents the cited-cited relation, which differs from the previous two as it does not directly relate the citations to the citing text: rather, this relation targets the links that are drawn by the citing writers between two or more sources external to the current text. It differs from function analysis in that it does not relate the citations to the citing text in order to understand their role in the current texts. Instead, the emphasis in this relation is on how two or more external sources relate to each other. In other words, the analysis in this respect investigates whether citations are consistent with one another: providing support or showing agreement between each other or whether the citations are presented in a way which shows them to be critical of each other or inconsistent with one another. The two cited-cited relations are citation-support and citation-engagement. It is important to note that support and engagement here are not synonymous with the citation functions as identified in this thesis. As citation functions, Support and Engage relate the cited proposition to the ongoing citing argument or

finding whereas the previous two relations refer to the relations identified by the writer but between citations, either to compare or contrast different sources.

Some of the citations analysed in Sections 4.2 and 4.3 in terms of stance and function can be reanalysed with respect to this relation. However, it is not possible to have a cited-cited relation without citing-cited (stance) and cited-citing (function) relations. Writers evaluate external sources in their texts to achieve certain functions, and occasionally draw links between the cited sources. Consider the following example which is represented in terms of the cited-cited relation in Figure 4.2. This example is analysed in terms of stance and function in the previous section (see the excerpt in Example 4.42).

Example 4.42 is revisited below to show how the citations relate to each other. This is an example of citation-support where one citation is shown by the writer to be consistent with another. The citing statements in this block are S2, S3, S4 and S5 and they include citations 1, 2, 3 and 4, respectively. Regardless of the various stance types that can be identified at the clause-level of each citation, all the citations are analysed and coded for citation functions and relations. This is in order to show how the two have been identified differently in the current study. The following figure (4.2) demonstrates this relation.[6]

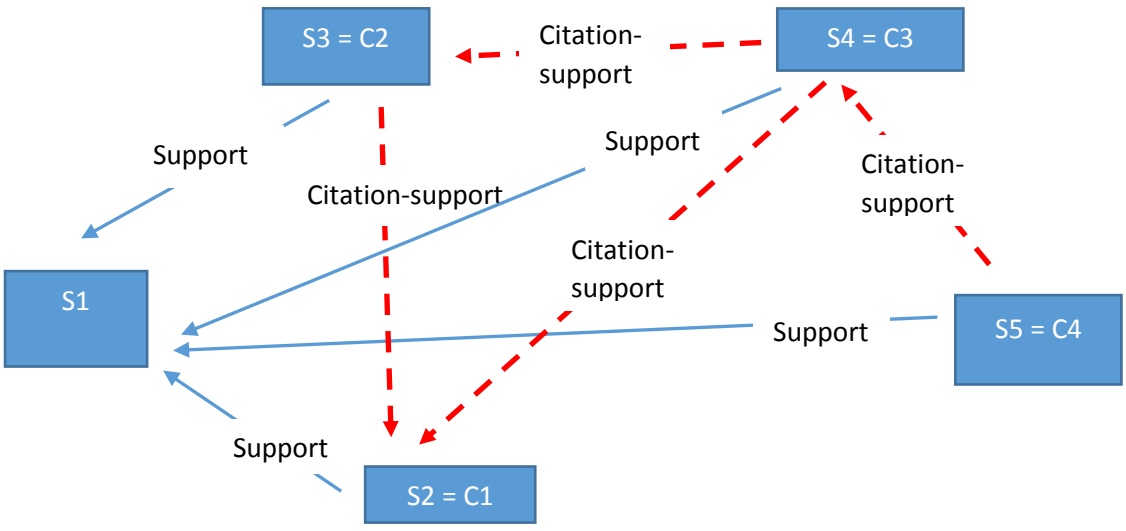

Figure 4.2: The cited-cited and cited-citing relations in Example 4.42 from Section 4.3

In Example 4.42, the first statement (S1) prospects the remainder of the paragraph. All the citations in this block Support the writer's main argument in S1 that the Marshall Institute Trio's work *has influenced top-level political decision makers' positions*. The first statement prospects further elaboration by the writer on the way in which the trio's work has been influential. At the cited-cited relation, each of the four citations (in S2-S5) supports its neighbouring citation, whether the preceding or following citation. The block is self-explanatory: the proposed information in the first statement (S1) is reiterated in the following statements (S2-S5) with cited evidence of such influence of the trio's work. All the citations show that the trio's work has been influential (even though the citing writer disagrees with some of the trio's work).

Below is another example of citation-support, see Figure 4.3 (this figure provides a visual representation of Example 4.38).

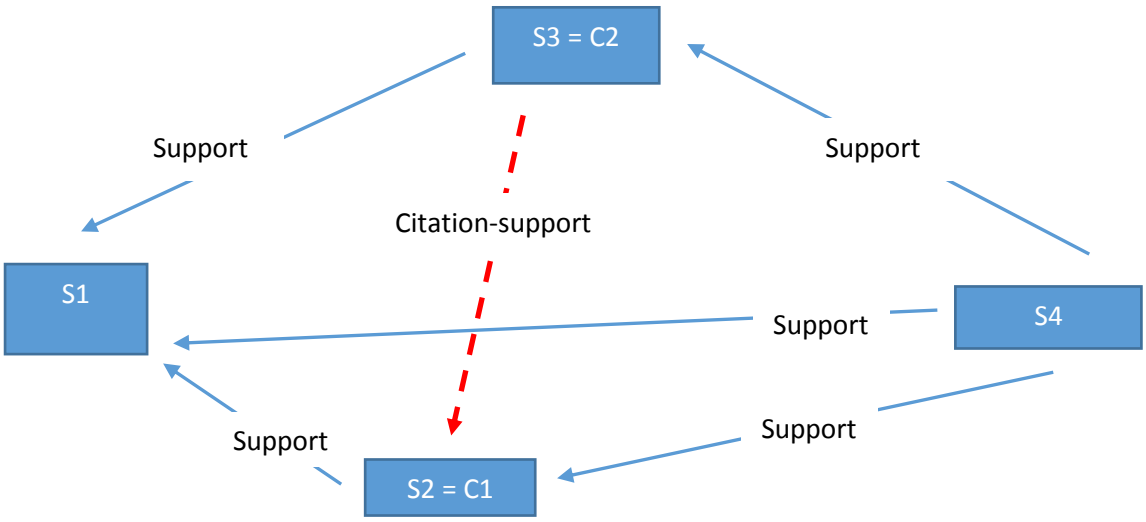

Figure 4.3: The cited-cited and cited-citing relations in Example 4.38 from Section 4.3

In Example 4.38, the links between the first and second citations (in S2 and S3) are less overt compared to those in Examples 4.42 and 4.43. The writer includes S3 in support of the other citation in S2. The mutual support between the citations becomes evident in the block. In other words, the writer avers the alarming increase of deforestation in S1, cite-in-Support one estimate in S2, *acknowledge*-in-Support the increase of the annual deforestation rate in S3 and finally, S4 (which is averred by the writer) re-evaluates and assumes as true the first three statements.

In addition to the fact that the cited information in the third statement (S3) Supports the information in the first and fourth statements concerning the increase of tropical deforestation, it also supports the cited information in the second statement (containing the first citation: C1). The second citation (in S3) shows the increase in annual deforestation in the period between 1981-1990 to reach *16.9 million hectares* compared to *11 million hectares* in 1980, as estimated in the first citation in S2. Also, the use of the verb phrase *has increased* in the third statement implies recognition and acceptance of the initial estimate in the second statement. The citing paper was published in 1990. The writer introduces the topic on increasing tropical deforestation in the preceding decade in the first statement. The subsequent two statements represent exemplary studies from that decade. They both boost the writer's argument in S1. The third statement (containing C2) also supports the first citation in S2. Both statements relate to the annual rate of deforestation and the latter shows the cited increase in that rate, which also supports the initial estimate of the former citation.

The second relation between the citations identified in this thesis is citation-engagement. This relation shows contradiction or inconsistency between two or more cited sources and is usually signalled by adversative discourse markers such as *however, nevertheless, on the other hand*. The following excerpt (Example 4.44) illustrates how this relation is identified at the level of the block.

Example 4.44:

> [1] *In Amazonia, the historical role of cattle ranching in deforestation is partly a result of the favourable tax and credit incentives received by cattle ranchers from the mid-1060s through mid-1980s which made converting forest to pasture more profitable than the sustainable use of already-cleared lands (REF).* [2] ***However****, even at the height of the government incentives programs in 1975 over 45% of all clearing in eastern Amazonia … was in large ranches that received no government subsidies (REF).* [3] In part, this reflects … [2009_McAlpine et al.]

Engagement between the sources is signalled by *However*. Both statements (or parts of them) are attributed to the references. They show two opposing arguments. The first attributes cattle ranching in deforestation partly to tax and credit incentives between the mid-1960s to the mid-1980s. On the other hand, the second statement shows that in 1975 over 45% of clearing was *in large ranches that received no government incentives*. It is important to stress that citation-engagement is not relevant to stance or to function; it shows how the sources interact with each other, not with the citing text.

The following figure (4.4) provides a visual representation of Example 4.43 whereby different citation relations are identified in the same block. The block demonstrates how writers can employ both cited-cited relation types to bring the sources into dialogue.[7]

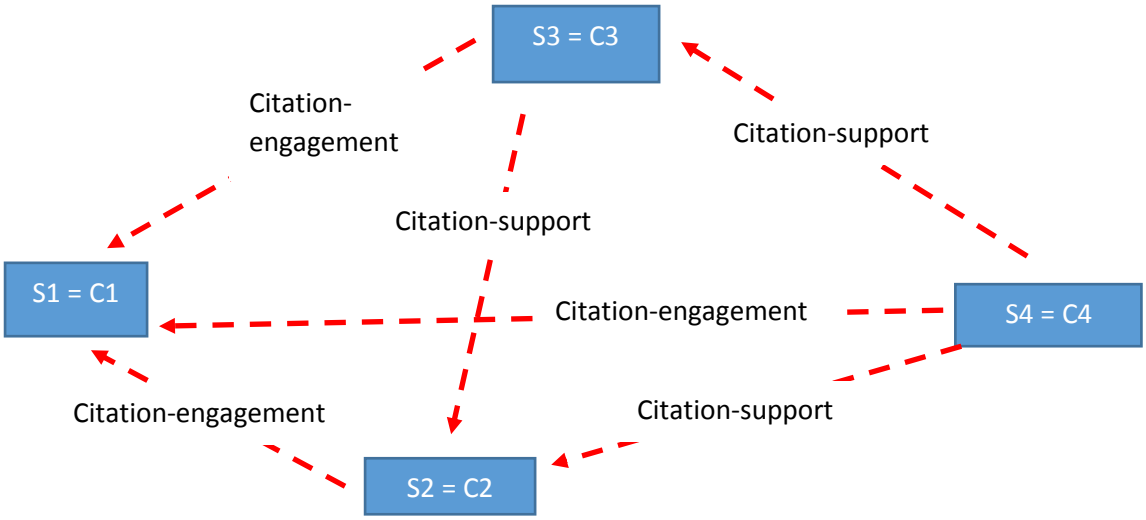

Figure 4.4: The cited-cited relation in Example 4.43

From the above figure, it is evident that the citations (in S2-S4) engage in critical dialogue with the first citation (S1). All the statements are attributed to sources. The first sign of citation-engagement is *But* at the beginning of the second statement (C2). This citation in turn (C2) is supported by the following two citations (C3 and C4). While the latter three citations show citation-engagement with the first citation, they also support each other. The third and fourth citations justify the claim in the second citation. In other words, they support the cited claim which appears in S2: that global warming has not yet been documented.

4.5 Summary

In Chapter 4, the proposed model and the categories have been exemplified and analysed with particular attention to the linguistic features related to the coding of citations in terms of form, stance and function. Analysis of the individual categories during the early stages of this research resulted in the identification of conventional and unconventional citation patterns. These patterns (or stance-function combinations), I would argue, provide a more comprehensive perspective of the investigation into how citations act, and are acted upon, in the citing texts. The introduction of the block as a new analytical unit has facilitated the identification of the conventional and

unconventional combinations in addition to the different relations identified between citations beyond the proposition-level of each citing statement.

The next chapter provides a quantitative overview of the results of the coding of citation form, stance and function categories. Chapter 5 presents the frequencies in the form of ratios. The ultimate focus is not on comparing the individual files (the twenty research articles): rather, the aim is to provide a quantitative comparison of the four sub-corpora to account for any diachronic or synchronic variation between the selected papers. In other words, the interest is in general tendencies and preferences amongst the selected papers to show synchronic and diachronic variation in citation practices. The quantitative results are presented in Chapter 5 and discussed with respect to the IDRD project in Chapter 7.

Notes

1- There has not been consensus on whether the structure *According to*, shows *acknowledgement* or *distancing*. For example, Martin and White (2005) treat it as an indicator of *acknowledgement* while Hyland (2005c) and Malmstöm (2007) perceive it as a marker of *distance*. I agree with the second view because it shows that what follows is from the perspective of the source alone. The surrounding context usually shows agreement or disagreement with the attributed information.

2- In the following examples, the statements which include citations are italicised and the linguistic signals of stance and/or function are given in bold.

3- There are many examples of highly reputable individuals, institutions and organisations in the selected papers such as FAO (Food and Agriculture Organisation), IPCC, UNEP (United Nations Environment Programme), and official records such as Tuvalu Census and GoT (Government of Tuvalu).

4- Examples such as 4.27 are uncommon in the data. Therefore, the clause- and proposition-levels are used interchangeably to refer to the analysis of form and stance at the citing-statement level. For this reason, implicit *contest* was identified in Example 4.27 in addition to *distancing* at the clause-level. It should also be noted that this thesis is concerned with the writer's, not author's, evaluation (i.e. *claims* not *refutes*, as in Example 4.27) in the identification of the writer's stance. However, the two positions are combined to give an indication of the writer's implicit evaluation of the source's argument.

5- Although some citations have been identified as multi-stanced and/or multi-functional, this section is reserved for the analysis of how citation function combines with explicit stance. As shown at the end of Section 4.2, implicit stance is usually confirmed by function analysis. In other words, if a citation is implicitly *endorsed* or *contested*, then the surrounding context tends to continue, and confirm, the implicit *endorsement* or *contest* (see Examples 4.17 and 4.27).

6- The cited-cited relation is represented by the red dotted arrows in Figures 4.2, 4.3 and 4.4 whereas the cited-citing relation (citation function) is represented by the blue arrows in Figures 4.2 and 4.3.

7- Figure 4.4 demonstrates the cited-cited relations between the first four citations in Example 4.43.

Chapter Five

Constellation and Time Analyses: a Quantitative Comparison

5.1 Introduction

The previous chapter has set out the model categories and has provided qualitative examples and a description of the way in which the data was coded, using the 10 analytical categories arising from the form, stance and function analyses. This chapter presents the quantitative results obtained from applying the analytical codes to the selected papers from constellations 1 and 5. The quantitative figures are then compared to the findings of the Interdisciplinary Research Discourse (IDRD) project, as reported in Thompson et al. (2017), see Chapter 3, Section 3.2. The findings will also be discussed in detail in Chapter 7.

The main aim of this thesis is to build, not to quantify, a model which is able to reflect the role of citation in interdisciplinary academic writing. The use of the block as a new unit of analysis in the study of citation functions makes it impossible to apply the model to all the papers in GEC, particularly when using a 'close-reading' method as in the current study. In this thesis, citations are examined in blocks in order to understand their roles in the texts. However, the qualitative analysis does raise a question regarding the quantitative application of the proposed model: whether the proposed categories are able to contribute to what we know about the differences between the constellations and time periods, as demonstrated in Thompson et al. (2017).

In the IDRD project (Thompson et al. 2017), all papers between 1990 and 2010 in the journal GEC were compiled into a corpus and divided into constellations using a version of Biber's (1988, 1989) multi-dimensional analysis (MDA) model (see Section 3.2 for more detail on the IDRD project). The analysis shows the first and fifth constellations to be the most "distinguished and contrasted" (Thompson et al. 2017: 18). Constellation 1 (namely: Quantification and Measuring) is described as "the most 'science like', reporting empirical work" (ibid, 21). Constellation 5 (Personal Voices), on the other hand, represents the social science-like component of the GEC corpus reporting social and political attitudes towards environmental issues.

Although the project did not investigate citation practices in GEC, it provides evidence of variation between the two constellations (the differences are summarised in Section

3.2). For example, Thompson et al. (2017) find that constellation 5 papers, unlike those of constellation 1, score highly on features associated with explicit argumentation. Explicit argumentation is realised in positively-loaded features such as modals, stance adverbs and conjuncts (ibid: 11). Similarly, the current thesis provides evidence for explicitness in the study of citation stance and function in constellation 5 papers. Consider the following two citations shown in italics in S2 and S3, which share the same block:

Example 5.1:

> [1] Negotiations about the initial distribution of permits will**, of course,** be dominated by the **self-interest** of participating nations.[2] ***Nevertheless, it is often maintained*** *that considerations of international* ***equity*** *will also play some part, even if only as providing 'focal variables', to use Barrett's term, around which the negotiations can be centred****..***[fn] [3] ***Indeed****, explicit* ***recognition*** *of the role of* ***'equity'*** *considerations* ***is contained in*** *the Framework Convention on Climate Change agreed at the 1992 Rio de Janeiro World Environmental Conference (INC 1992).*[fn]
>
> [4] **However**, the widely **conflicting interests** of different countries as regards the extent and form of measures to restrict carbon emissions **makes it unlikely that genuine equity considerations** will play **any role** at all as compared with naked bargaining, as the March/April 1995 Berlin conference on the subject has demonstrated. [5] **Furthermore**, as will be argued below, **it is highly unlikely anyway that any widely accepted ethical principles could be found that have any bearing on the equitable international allocation of tradable emission permits**, in spite of the impression to the contrary given by discussion of various 'equitable' allocations of emission permits. [1995_Beckerman and Pasek]

Form, stance and function analyses of this example appear in Chapter 6 (Example 6.8). This example is introduced here to show the difference between explicitness as understood in the IDRD project and the current thesis. In the project, explicit argumentation is associated with the concern for "guiding the reader's interpretation" (Thompson et al. 2017: 11) of the writer's stance. Stance according to that view is not limited to the writers' evaluation of the cited information. Rather, it is oriented towards the writers' stance towards the topic under discussion and the way in which the writers choose to draw links between propositions. The project focuses on stance conjuncts, because it uses Biber's methodology, such as *Nevertheless*, *However* and *But* as signals of explicit connections between propositions. However, that is not to imply that their analysis excludes other features associated with the writers' stance such as stance adjectives (e.g. *conflicting* and *unlikely*).

In the current thesis, the ultimate focus is on the citations themselves (in S2 and S3). The linguistic resources in the surrounding context (the block) are examined to understand the role of the citation in the block (i.e. its function in the citing block). Form and stance are analysed at the clause-level of every citation. Stance analysis is limited to the writers' evaluation of the cited information at the sentence-level. The other sentences (S1, S4 and S5) are included to study the linguistic resources that signal the citation function. In summary, I look at signals of citation/attribution such as *it is often maintained that* and *is contained in*, signals of topic (e.g. *self-interest*, *equity* and *distribution of permits*) and signals of contrast/change in the argument (e.g. *Nevertheless, However* and *But*) as well as those signalling continuation of the argument (e.g. *Indeed* and *recognition*, for more of those signals see Thompson and Zhou 2000 on disjuncts and concessive relations; and Hyland 2005b, c). This example is discussed in detail in Chapter 6.

Overall, explicit argumentation in the IDRD project is concerned with stance at the whole text-level. On the other hand, explicit argumentation is approached from a different perspective in the current study. Stance analysis is conducted at the clause-level and thus only the features appearing in the citing statement are highlighted in the study of the writers' stance. In the study of citation functions, the stance conjuncts (e.g. *Nevertheless*) are examined at the block-level to understand how the citation functions and relates to the citing text. In other words, these markers could be seen as indicators of the writers' stance beyond the clause-level of the citing statement, which is taken as an indication of the citation function (see Chapters 3, 4 and 6 on the identification of stance and function).

Despite terminological and methodological differences, the findings from the current study are consistent with those in the IDRD project. For example, the block analysis of citation functions in this thesis shows evidence of explicit and implicit links between propositions. Those links not only act to show the writers' stance towards the topic or cited propositions but also to signal how the citations function in the block. The signals of contrast (e.g. *However* in S4) show that the writers are involved in critical dialogue with the sources' views, thus the function is to Engage with the two citations.

The current thesis provides evidence (in this and the following chapters) of synchronic and diachronic variation between the four sets of data representing the two

constellations and time periods as discussed in Section 3.3. The four sub-corpora (or the data sets) comprise:

- Five papers from the early period (1990-95) of constellation 1
- Five papers from the later period (2008-10) of constellation 1
- Five papers from the early period (1990-95) of constellation 5
- Five papers from the later period (2008-10) of constellation 5

The current chapter provides quantitative evidence of variation, even though the reliability of the evidence is reduced by the small number of papers in the current thesis, due to the difficulty of applying the multi-level analyses of citations to more than the 1186 citations covered by this thesis. The results therefore reflect tendencies in the selected papers and preferences for some citation forms, stance types and functions. The current thesis uses a text-based rather than corpus-based methodology and the results are included in order to compare the findings with those obtained in another corpus-study (as reported in the IDRD project, see Thompson et al. 2017).

Overall, the main findings from the quantitative analysis of the four sub-corpora can be summarised in the following two points: the first point is that, in the early years, there were fewer citations. This is possibly because the journal (GEC) was new in 1990 and building a new field and hence there was possibly less to refer to. In the recent papers, the use of citations increased by more than 130%. For example, in constellation 1, the total number of citations increased from 215 in the selected papers between 1990 and 1995 to 500 citations in the more recent papers representing the period 2008-10 (a 132% increase in citation frequency). In constellation 5, the observed increase is from 140 to 330 citations, comprising a 135% increase in citations.

The second and more important finding is that in the first five years of the journal (as shown by the 10 papers from constellations 1 and 5 in that period), constellations 1 and 5 were quite different from one another in terms of their citations, whereas in recent years they are closer to each other (as shown by the 10 papers representing the period 2008-10 in both constellations). This is exhibited in the distribution and proportions of form, stance and function categories in the four sub-corpora; see Figures 5.1, 5.6 and 5.11. The following three sections provide quantitative evidence of synchronic and diachronic variation between the four sets of data in terms of citation form, stance and function, respectively. Then, Sections 5.5 and 5.6 show the results of different, yet

relevant, analyses. Section 5.7 summarises the main findings and argument of this chapter.

5.2 Citation forms

Figure 5.1 shows the averages of integral and non-integral citations in the constellations and time periods. Similarly in the next two sections (5.3 and 5.4), Figures 5.6 and 5.11 provide aggregate views of the distribution and spread of the stance and function categories in the four data sets. Figure 5.1 will be discussed at the end of this section, but meanwhile we turn our attention to the individual papers. Figures 5.2, 5.3, 5.4 and 5.5 show the ratios of integral to non-integral forms in each of the 20 papers. The reason is that such representation exhibits the internal variation within each sub-corpus. Each of the form, stance and function sections in this chapter presents the ratios which represent the category proportions in each of the papers. However, the ultimate focus of this chapter is not on the internal variation between the papers; instead, it is on the variation between the four sets of data.

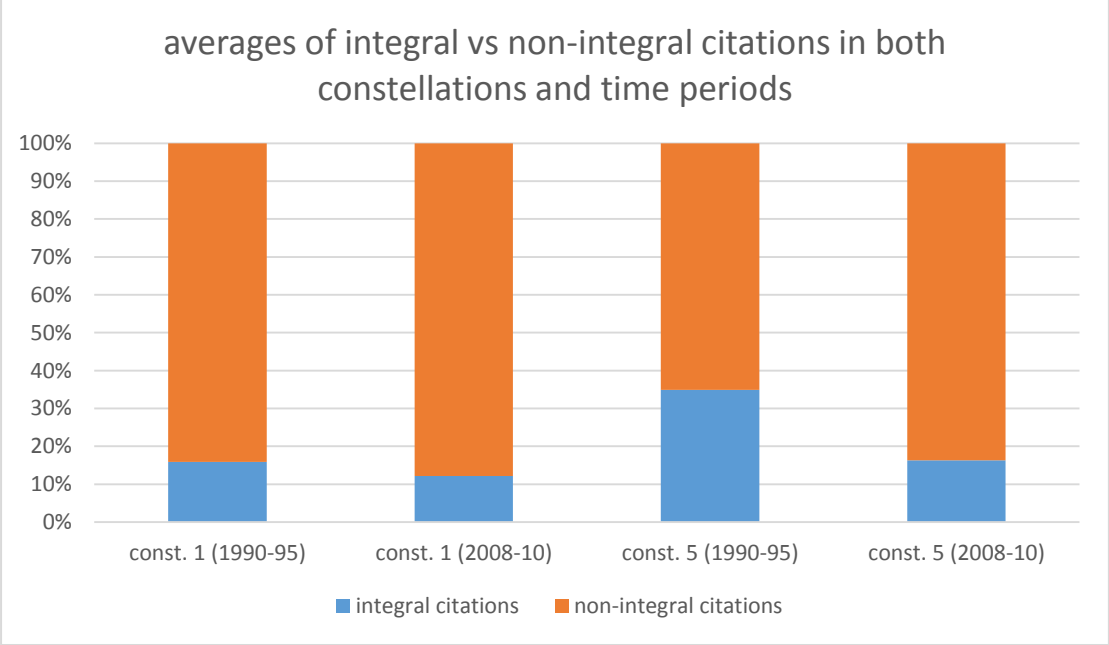

Figure. 5.1: Ratios of integral vs non-integral citations in both constellations and time periods

Evidently, there is a preference for non-integral forms in the selected papers (except for one paper, Beckerman and Pasek 1995, see Figure 5.4 and Chapter 6). While the literature tends to focus on integral forms in the analysis of citations (e.g. Pickard 1995; Harwood 2009; Manan and Noor 2015), the data shows a preference for non-integral forms in most of the papers.

Figure 5.1 shows a diachronic increase in the use of non-integral forms in both constellations. These forms rose from an average of 84.1% between 1990-95 to 87.8% between 2008-10 in constellation 1.[1] As noted in Chapter 3, constellation 1 represents the scientific component of the GEC journal (Thompson et al. 2017). The high proportion of non-integral citations in this constellation (see Figure 5.1) is consistent with the emphasis given to the message rather than to the individuals or organisations that produced it. Thompson (2001) notes that in a scientific subject such as Agricultural Botany, PhD writers prefer non-integral forms when citing external work. He attributes this preference to a concern with the cited message rather than with its producer. Similarly, I argue that the dominance of non-integral forms in both constellations could be attributed to the writers' concern with the message rather than the authors (see Example 4.13).

With respect to constellation 1, this finding is consistent with those reported in the IDRD project that papers in this science-like constellation are characterised by a "de-humanised approach" and "implicit argumentation" (Thompson et al. 2017: 19). At the clause-level of the citing statement, integral forms tend to position the source's name in subject position (e.g. *Campbell (2000) argues*). This act clearly gives prominence to the source's contribution over their argument. Also, integral forms usually contain reporting verbs which clearly signal the writer's stance towards the cited authors (e.g. *demonstrates, overlooks, claims*). Therefore, integral forms seem to be associated with a humanised approach and explicit argumentation at the clause-level of the citing statement. The non-integral forms however, usually do not include the above explicit markers as the writers usually bring to the foreground the cited arguments and consign the sources to the background. Non-integral citing could be seen as a sign of preference for the physical world over the world of ideas (ibid: 19) in that writers focus on the activities related to environmental issues more than the perspectives portrayed by the source (e.g. *The impact of deforestation … is unprecedented (REF)*). In this example, there is clearly a concern with deforestation (as an environment-related issue), rather than with the source of the argument, who appears between parentheses. Thompson et al. (2017: 11) treat such cases of "absence of explicit agency" as markers of implicit argumentation. Along these lines, I draw a distinction between integral forms whereby the author's name plays a role in the construction of the citing statement and non-

integral forms as in the above example where the author's name plays no grammatical role in the construction of the citing statement.

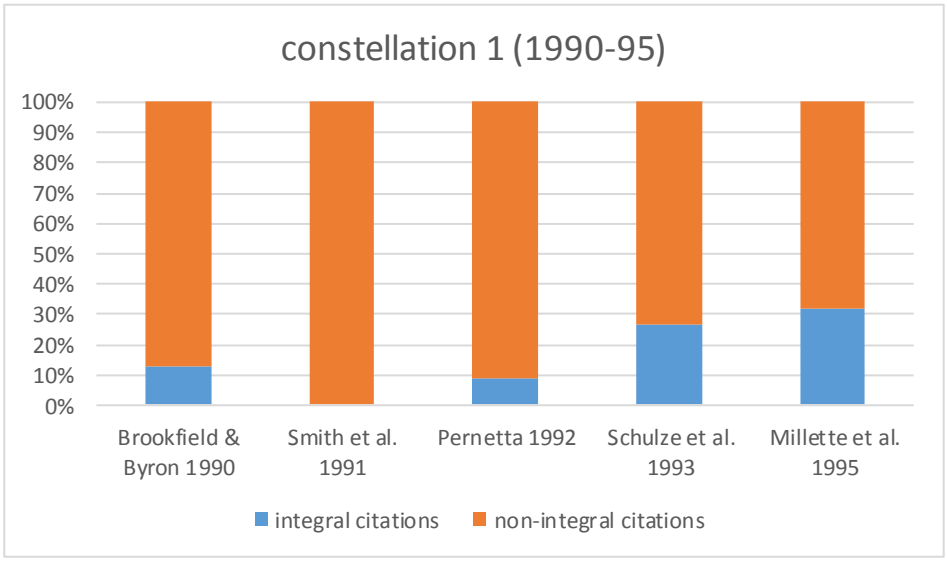

Figure 5.2: Ratios of integral vs non-integral citations in the early papers of constellation 1

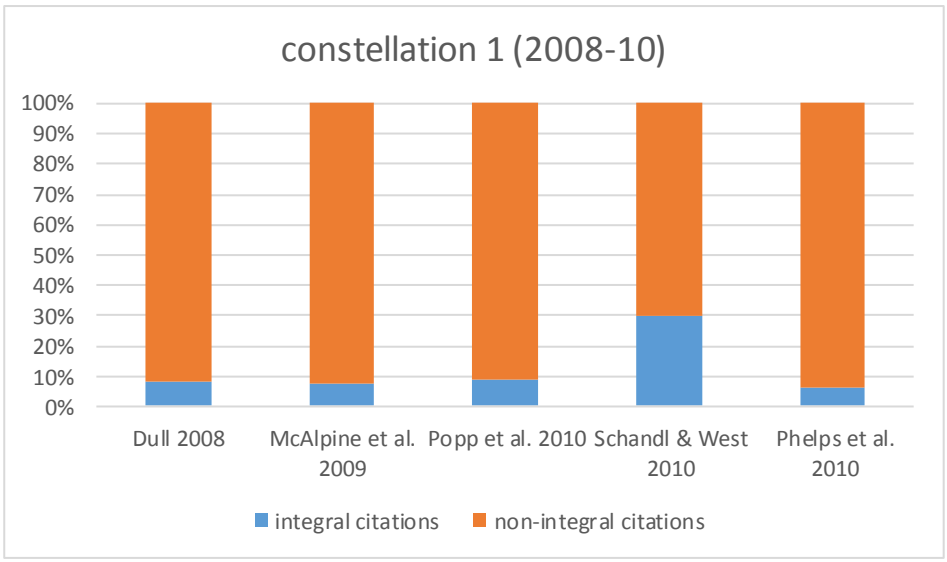

Figure 5.3: Ratios of integral vs non-integral citations in the recent papers of constellation 1

In Thompson et al. (2017), explicit and implicit argumentation are linked to the writer's "concern, or lack of it, for guiding the reader's interpretation" (ibid: 11) of the writer's stance and/or the links drawn between propositions. The use of stance markers and conjuncts for example, indicates to the reader, the writer's stance and how the propositions relate to each other (e.g. in the use of *However* which shows a change in the direction of the argument). I would argue that some of these markers can be

observed in the citing statement. For example, it is common for writers to indicate their stance in citation explicitly via the use of reporting verbs or evaluative lexis (e.g. *Indeed, compelling, admittedly, but*). These explicit markers can be observed in the citing statement, irrespective of the block. Also, the absence of such markers could be seen as a sign of implicit argumentation at the level of citing statement. In this view, I would argue that while Thompson et al. relate markers of explicit and implicit argumentation to the links drawn "between propositions, and the author's stance" (ibid: 11), such markers can also be observed at the level of citing proposition alone. Of course, the citation block makes the identification of such markers more useful for the understanding of the writers' stance and the links they draw between propositions, which are considered in function analysis (see Example 5.1 above: e.g. *Nevertheless, Indeed, However, Furthermore* in S2, S3, S4 and S5, respectively). Yet, in this section, it is argued that the increase of non-integral forms can be linked to the characteristics of constellation 1, mainly to the "de-humanised approach" and implicit argumentation as found in Thompson et al. (2017). The higher proportion of non-integral citations in this constellation indicates a less apparent stance (as opposed to integral forms which make the expression of stance easier as in *Smith (2000) claims*; refer to Chapter 4 for more examples and Section 5.5 below for a summary of the frequency of implicit stance in the four sets of data).

Turning to constellation 5, in the early years of the journal constellation 5 was quite different from constellation 1 in terms of the frequency of integral citations (see Figure 5.1). Figure 5.4 shows the internal variation between the papers in the first five years of constellation 5. Compared to the other three data sets which represent the other constellation and/or time periods, the selected papers from the early period of constellation 5 notably contain higher proportions of integral citations. In fact, one paper from this period (Beckerman and Pasek 1995) contains more integral than non-integral citations (examples from this paper are discussed in Chapter 6).

Unlike constellation 1, papers in this constellation have been found, by Thompson et al. (2017: 20) to be associated with explicit argumentation and a concern "with human perspectives on the environment … and also with social perspectives of science". Although the authors do not link these two features to citation, their findings are consistent with the findings of the current study. A clear example of these two features may be seen in Example 5.1 above. In the above example, the writers make the steps of

argumentation clear to the reader. In other words, they first present their view in S1 about the *Negotiations about the initial distribution of permits*. Then, starting with *Nevertheless, it is often maintained that* in S2, they present an opposing view which they attribute to the cited sources. They continue this view in S3 (*Indeed, explicit recognition of … is contained in*) and then Engage with the cited views in S4 (*However, the widely conflicting interests … makes it unlikely that*).

What is remarkable in S3 is that the noun *recognition* (as in *explicit recognition of the role*) indicates the writers' concern with the human perspectives on the environment and also with the social perspectives on environment-related issues, as reported by Thompson et al. (2017). Therefore, the current study provides additional evidence of some features reported in the IDRD project, but in terms of citation practices in the journal. This concern with the human perspectives could well be the reason why, in the selected constellation 5 papers, the ratio of integral citations is higher than those in the constellation 1 papers. The cited authors' names are usually brought to the forefront when the writer wishes to emphasise their contribution, argument, or perspective and/or how their arguments relate to the ongoing discussion. Consider the following two examples:

Example 5.2:

> ***O'Neill goes as far as to say that****, as regards international justice, utilitarian consequentialism is a non-starter.* [1995_Beckerman and Pasek]

Example 5.3:

> *What about contractarian theories of justice,* ***such as their recent formulation by Rawls****,* [1995_Beckerman and Pasek]

Examples 5.2 and 5.3 include integral citations. In these examples, the sources' contributions are highlighted in some way. For example, in 5.2, the cited author (O'Neill) is integrated into the citing statement in a way that shows that the source (O'Neill) extends critical views of utilitarian theory of international justice expressed in the preceding statement. The writers in the preceding statement argue that utilitarian theory has its own problems and O'Neill is reported as reformulating this statement by saying it is a *non-starter*. Also, more importantly, the phrase *goes as far as to say that* evaluates O'Neill's contribution as counter to expectation. The effect of this reporting phrase is difficult to replicate in a non-integral citation, unlike other reporting structures

(e.g. *O'Neill says that* or *According to O'Neill*) which can easily be paraphrased in non-integral citations. The choice of the integral citation form rather than the non-integral is determined by the inclusion of the explicit evaluative meaning (i.e. it is determined by the writers' choice to evaluate O'Neill's contribution in a way which cannot be easily achieved using non-integral forms of citing).

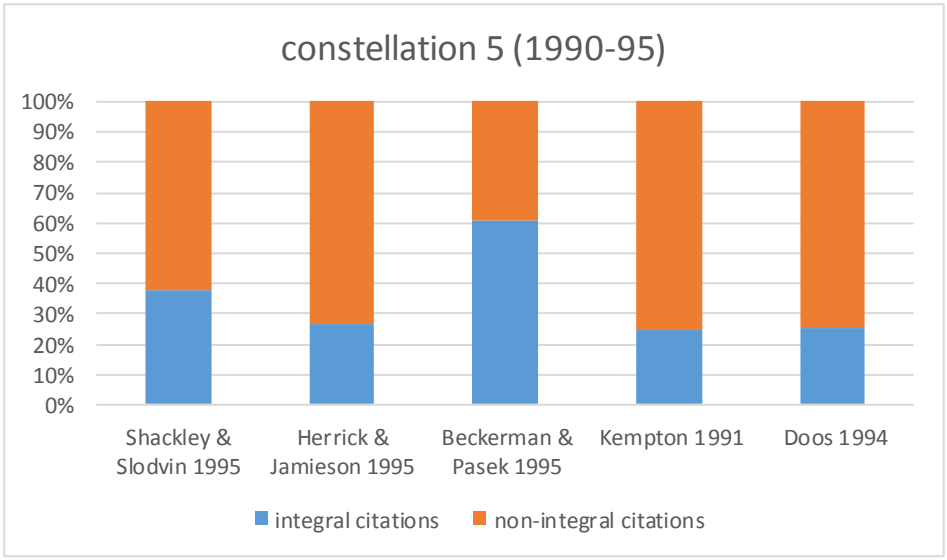

Figure 5.4: Ratios of integral vs non-integral citations in the early papers of constellation 5

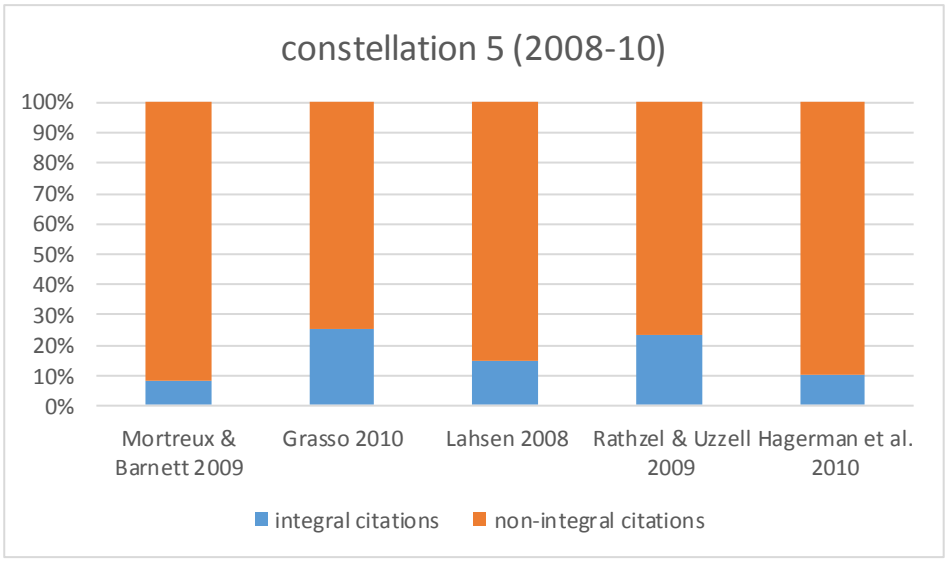

Figure 5.5: Ratios of integral vs non-integral citations in the recent papers of constellation 5

In Example 5.3, the writers *acknowledge* the role of Rawls in the recent formulation of contractarian theories of justice. The writers could have cited Rawls non-integrally (so

that it plays no grammatical role in the construction of the citing statement) but instead they stressed his contribution to contractarian theory. Overall, the two examples show some of the reasons why writers may wish to foreground authors' names or contributions. Integral forms usually require stance to be expressed more explicitly than do non-integral forms, as in the above examples.

Figure 5.1 shows the averages of integral and non-integral citations in the four data sets allowing a comparison between constellations and time periods. It is interesting that in the early years of constellation 5, the proportions of integral to non-integral citations were quite different in comparison to those in the papers in the same period of constellation 1, but that the more recent papers in constellation 5 converge with the recent papers in constellation 1. In other words, the average of non-integral citations in the recent years of constellation 5 reverts to 83.73%, a percentage which is similar to the averages of non-integral forms in both time periods of constellation 1, which has slightly changed over time in terms of citation forms.

The results reveal how the change with regards to citation form has occurred mainly and more notably in constellation 5. This constellation shows more noticeable diachronic variation in the proportions of integral to non-integral forms. This raises the inevitable question about why constellation 5 has changed more than 1. There are at least three possibilities: 1) that it was an editorial decision to use non-integral forms more frequently, 2) that it is the inevitable consequence of the journal being more focused and the discipline becoming established (i.e. it might be the case that constellation 5, Personal Voices, was a new kind of research in GEC in the early 1990s when the journal started and 20 years later this constellation was no longer the same and the citation practices converged with other constellations) or 3) the most likely reason I believe, is that it was an issue of disciplinary style. In other words, writers in the early years were establishing the field which could have resulted in a relative increase of integral forms in comparison with the other period and to constellation 1. Integral forms might have been chosen because they facilitated engagement with sources and as the discipline became more focused, writers began to adopt a more scientific citing style. In this regard, it is also possible that writers in recent years realised that their audience increasingly included environmental scientists. Therefore, they changed their citation practices to match those of scientists (as in constellation 1). This writer-awareness issue is a potential cause of the recent drop in integral forms and

rise in non-integral forms in both constellations. In general, the increasing preference for non-integral forms allows writers to manipulate the use of citations (e.g. in passing more implicit stance, see Section 5.5).

5.3 Stance in citation

Figure 5.6 provides an aggregate view of the distribution and proportions of the stance categories in the four data sets. Then, Figures 5.7 and 5.8 showcase the distribution and proportions of stance types in all the papers from constellation 1. It is evident that the dominant stance type in the two time periods is *acknowledge*. In fact, *acknowledge* is the most common stance type identified in all the papers. This finding is also consistent with research reported in Pickard (1995): that writers most commonly cite previous research to show they have read extensively within the field of study. The second most common stance type in constellation 1, immediately following *acknowledge*, is *endorse*. This could well be attributed to the identification of implicit *endorsement* which tends to be realised on the surface of explicit *acknowledgement* (see Section 4.2 in Chapter 4). In both time periods of constellation 1 (see Figure 5.6), *acknowledge* accounts for more than half of all the identified stance types, followed by *endorse*, with *endorsing* citations comprising more than 40% of the stance types. This relatively high frequency of *endorsement* is partly a result of the double-coding, as demonstrated in Section 4.2 under implicit stance or 'multi-stanced citations'. In such cases, both explicit and implicit stances are counted which justifies the increase of both stance types in this constellation. For example, when writers *acknowledge* statements that are construed as facts, they also implicitly *endorse* them as the *acknowledgement* becomes contractive in this sense (refer to *acknowledgement* of facts versus *acknowledgement* of assessments in Chapter 4). In such cases, both the explicit and implicit stances are counted in the analysis. This multiple-coding resulted in the increase of *endorsement*, which came second in rank, following *acknowledgement*. However, this by no means implies that explicit *acknowledgement* always co-occurs with another stance type, such as implicit *endorsement*. Rather, it simply means that while there are instances of *acknowledgement* (e.g. *Smith argues*) and instances of *endorsement* (e.g. *Smith demonstrates*), there are also instances in which implicit stance may be identified in addition to the explicit stance.

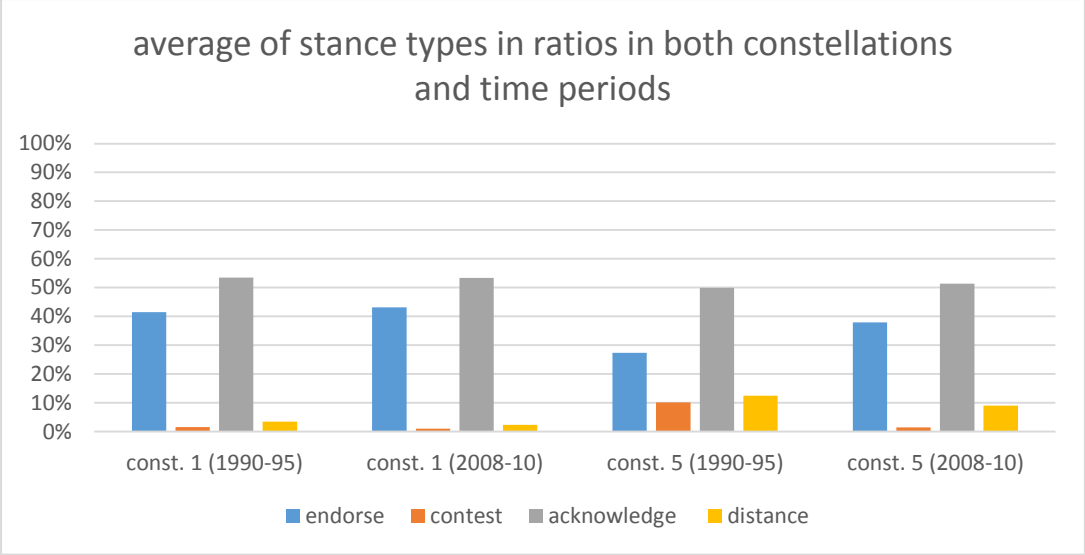

Figure 5.6: Stance type averages (in ratios) in both constellations and time periods

In constellation 1 (see Figure 5.6), *distancing* citations come third and finally *contest* is the least used type. The latter two types are relatively minimal in terms of their occurrence in this constellation. This might suggest that the writers in this constellation, Quantification and Measuring, do not negatively engage with sources, at least at the clause-level, as often as in the other constellation (also noted in Thompson et al. 2017). This constellation has maintained its relative distribution of stance types over time (as Figure 5.6 shows).

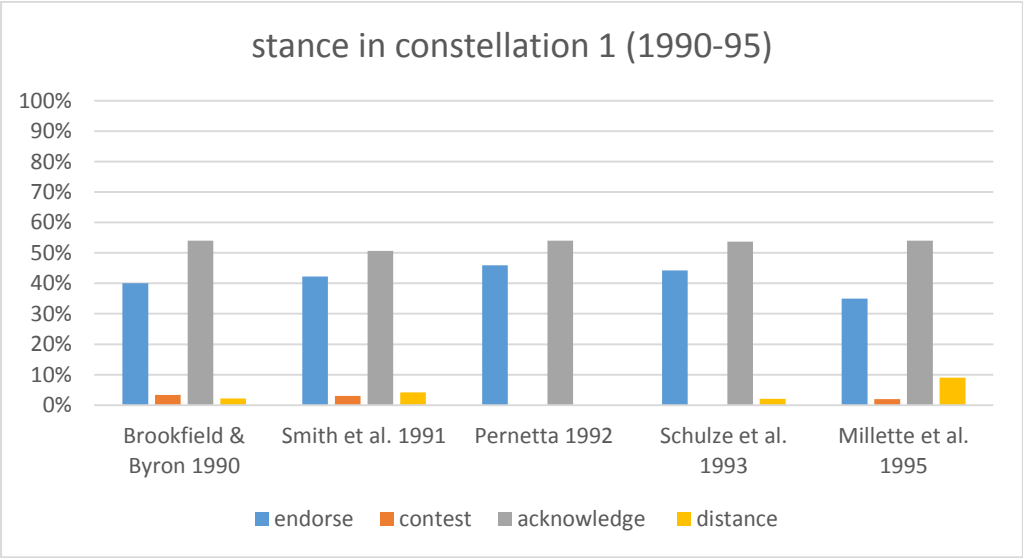

Figure 5.7: Stance type ratios in the early papers of constellation 1

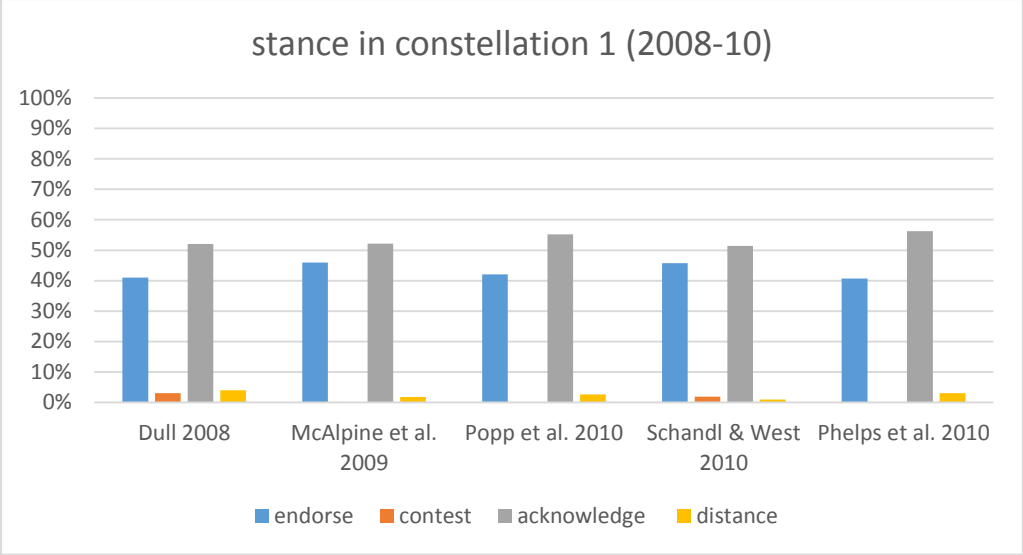

Figure 5.8: Stance type ratios in the recent papers of constellation 1

Acknowledge, followed by *endorse,* also takes the lead in constellation 5 (see Figures 5.6, 5.9 and 5.10). However, the proportions and distributions of the other stance types are different in this constellation. Figure 5.9 displays that there is internal variation between the papers in this time period of constellation 5. In other words, there is a more noticeable spread of stance types in this time period. While in constellation 1 *distance* and *contest* were minimal in terms of their occurrence, the case is different in constellation 5, which shows higher proportions of these two stance types, at least in the early sub-corpus. For the more recent papers however, Figure 5.10 shows a relative convergence with both time periods of constellation 1. Figures 5.9 and 5.10 show the internal variation in constellation 5 papers, and also demonstrate that the more recent papers in this constellation show relatively similar proportions to the constellation 1 papers. In other words, the proportions of *distance* and *contest* decrease in favour of *endorsement* (see Figure 5.6), although the decrease is more noticeable in *contest* than in *distance*. Figure 5.6 summarises the synchronic and diachronic variation between the two constellations and shows the diachronic increase of *acknowledgement* followed by *endorsement* in constellation 5 to converge with constellation 1.

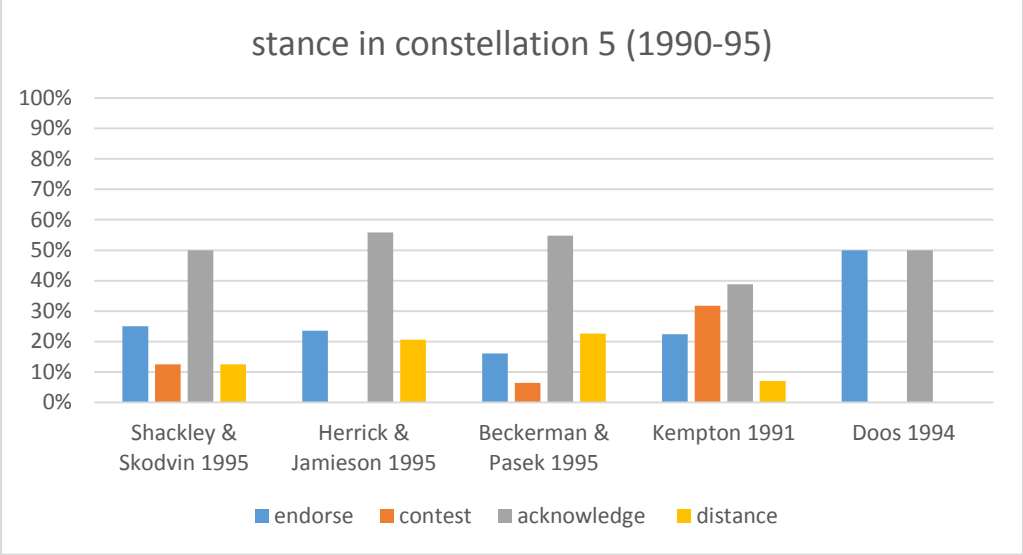

Figure 5.9: Stance type ratios in the early papers of constellation 5

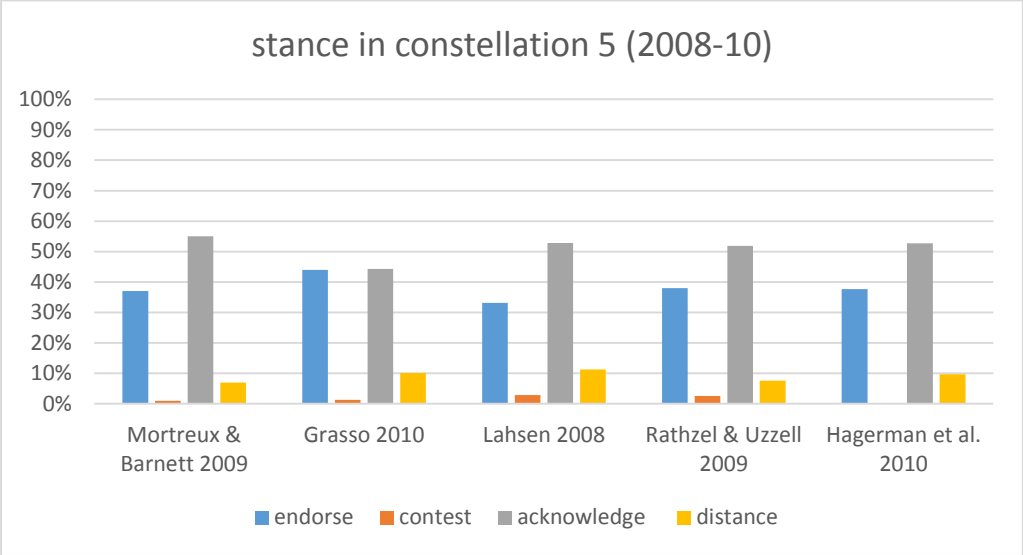

Figure 5.10: Stance type ratios in the recent papers of constellation 5

Up to this point, the clause-level analyses of the citation form and stance have shown a clear difference between the early years of constellation 5 and constellation 1. Form and stance analyses have shown divergence in the early years and later a convergence in the more recent years. The difference was apparent in the early years of constellation 5. The most obvious difference is the increase of *distancing* and *contesting* proportions at the expense of *endorsing* citations.

While the occurrence of *distancing* citations remains relatively high, in both time periods of constellation 5 compared to constellation 1, *contesting* citations were only higher in some papers of constellation 1 (e.g. Kempton 1991). This also might be

attributed to explicit engagement with the cited sources. *Contesting* citations exhibit contradiction with, or criticism of, the sources, particularly in the early papers. This might be the case when the field was being established and other authoritative voices and opinions were being questioned or challenged. Overall, *contesting* citations were unexpectedly low in both constellations compared to other stance types, such as *endorsement*. However, this finding is consistent with Pickard's (1993) finding that only 5% of citations in her data involved negative evaluation of previous work. This might suggest that engagement with the sources is deferred to subsequent statements (at the block-level). This possibility is revisited in function analysis (Section 5.4) and also in the following two chapters.

Acknowledgement and *endorsement* have increased over time in constellation 5, and although *distancing* has correspondingly decreased in the same constellation, its ratios remain considerably higher than in constellation 1. *Contesting* citations were relatively higher in the early period of constellation 5 (see Figure 5.6), subsequently they maintained the same percentages as constellation 1. Generally, the recent convergence between the two constellations at the clause-level may indicate that the journal has become focused in recent years as it bridged the gap between the two constellations with respect to citations. In summary, there seems to be more evaluation of citations in constellation 5. Yet, the two most frequent stance types are still *acknowledgement* and then *endorsement* in both constellations and time periods.

5.4 Citation functions

The first figure in this section (Figure 5.11) provides an aggregate view of synchronic and diachronic variation in the occurrence and distribution of the citation functions in the four data sets. Figures 5.12 and 5.13 show the ratios of citation functions in constellation 1. Perhaps not surprisingly, the dominant function is Support in all the papers representing the two time periods. In scientific papers, Support is a common function of citing. Scientific writers usually present their research as a continuation of existing knowledge in a given field and cite previous related research to provide theoretical background for the adoption of a specific method/approach, to confirm previous findings or to justify their own findings (Hyland 1999, 2000, 2009).

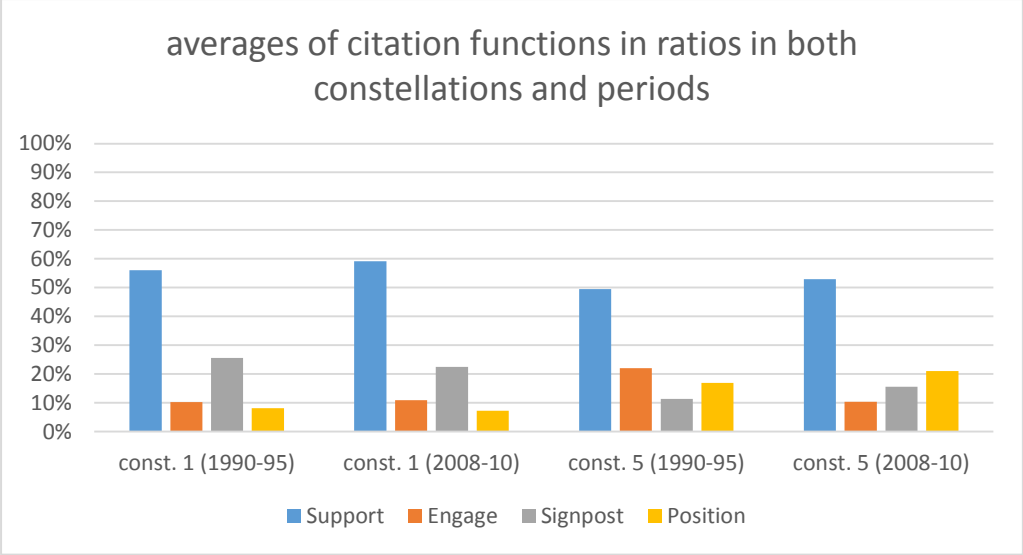

Figure. 5.11: Averages of citation function ratios in both periods and constellations

If all the citations in this constellation were conventional (i.e. showing a match between the identified stance and function types), one would expect the function proportions in Figures 5.12 and 5.13 to match the stance ratios in Figures 5.7 and 5.8. For example, in the first paper of Figure 5.7 (by Brookfield and Byron 1990), *endorsement* accounted for 40% of the stance types identified in the paper while Support, the corresponding function, accounted for 60% of the identified functions in the same paper as in Figure 5.12. This clearly shows that stance and function are different yet related aspects in citation analysis. It also shows that evaluation is not a function of citation (as Petrić 2007 and Dontcheva-Navratilova 2016 tend to argue). In other words, writers evaluate sources or cited propositions to achieve one or more functions. The stance-function combinations are extremely useful in the understanding of citation roles in texts and how writers prefer certain patterns of form, stance and function when citing.

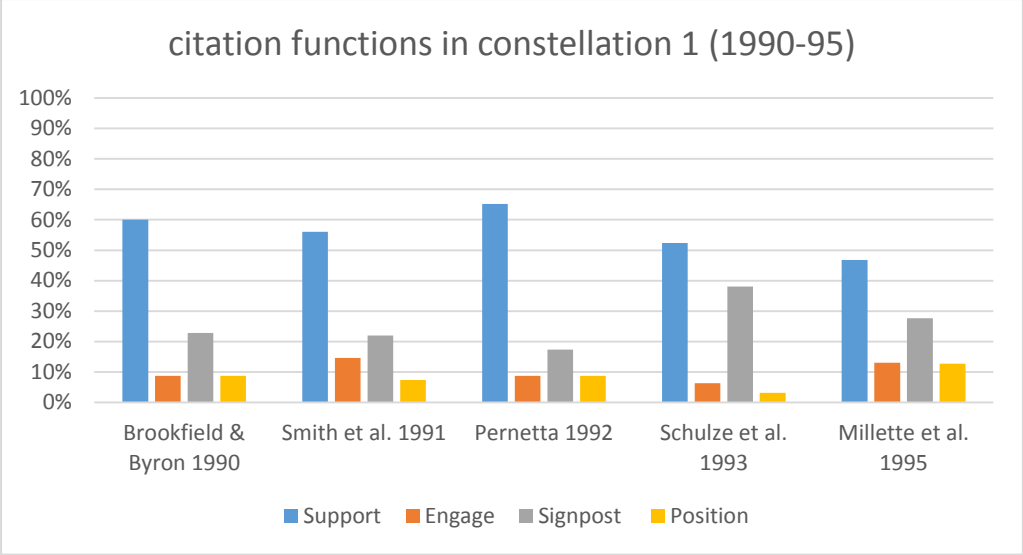

Figure 5.12: Citation function ratios in the early papers of constellation 1

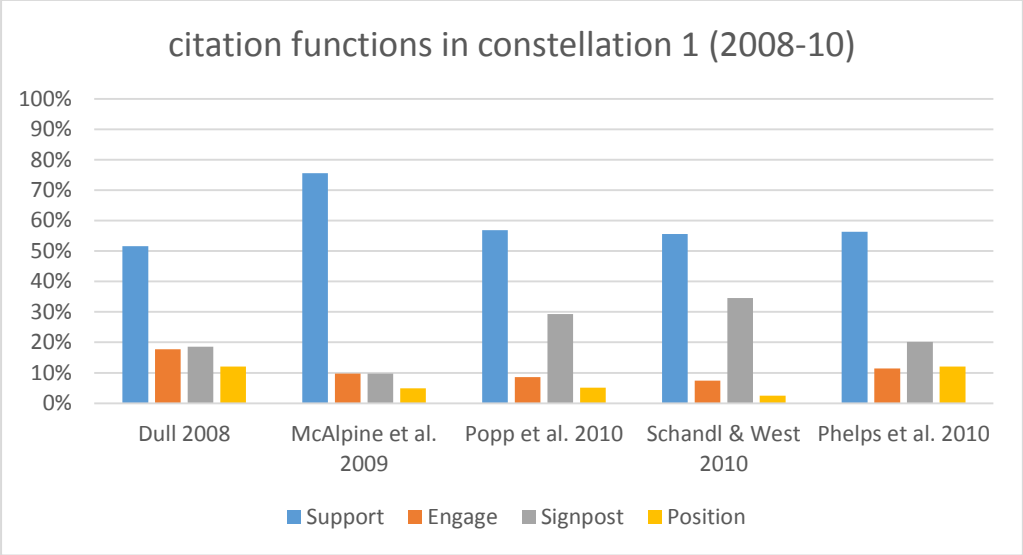

Figure 5.13: Citation function ratios in the recent papers of constellation 1

The ratios of function types in Figures 5.12 and 5.13 are different from the stance ratios in Figures 5.7 and 5.8, which represent stance ratios for the same papers. This thesis has clearly shown variation between stance proportions on one side and the corresponding functions on the other. For example, while *contest* and *distance* were both minimal in constellation 1, function analysis shows higher proportions of the corresponding functions Engage and Position for the same citations appearing in the constellation 1 papers. This clearly shows the importance of identifying the unconventional combinations. In other words, as noted in Chapter 4 (Section 4.3), if all the citations were conventional, then comparing the stance and function proportions

would be a mechanical exercise. This finding also shows that writers in constellation 1 are not putting themselves in conflict with cited materials at the clause-level, which is the reason behind the low occurrences of *contest* and *distance*. When writers choose to go into critical dialogue with the cited authors in constellation 1, they do that in a way that is not apparent at the clause-level (i.e. in the citation block, which is picked up in function analysis). In fact, this thesis finds that even on occasions where writers put themselves in conflict with cited views, they prefer to do so at the block-level, as in one paper from constellation 5 (see Chapter 6 for detailed discussion). In that paper, the writers construe a research environment to which they are antagonistic, however, that paper only contains two instances of *contest*. On most occasions, the writers *acknowledge* a source or an argument at the clause-level and Engage, in critical dialogue, with the source or their argument.

Acknowledge is identified as the most frequent stance type (see Figure 5.6), so one might expect the corresponding function of Signpost to be most frequently observed. However, since there are unconventional combinations (e.g. *acknowledge*-to-Support), function analysis has revealed some unexpected patterns. A given citation may show a certain stance type (e.g. *acknowledge* at the clause-level) while also performing a mismatching function at the block-level (e.g. Engage). In other words, it might seem at face value from stance analysis that a writer is merely *acknowledging* another source, and the corresponding function may thus be assumed to be Signpost. Yet the block-level analysis of functions can, and usually does, yield a different function. In constellation 1, the most frequent function is Support in both time periods. The dominance of *acknowledge* and Support in this constellation suggests that the writers use the *acknowledged* cited materials in Support of their ongoing arguments beyond the proposition-level of the citing statement, more than Signposting the readers to external research.

Following Support, Signpost is the second most frequent function in constellation 1. It is notable in constellation 1 that the majority of citations either provide Support to an argument or finding in the block or they are neutral, serving to guide readers to external sources. Little space is thus left for the negative function Engage, or the *distancing* function Position. This is largely due to the nature of this constellation, Quantification and Measuring, which is less engaged with the world of ideas (Thompson et al. 2017). The cited research mostly sets the scene for, or justifies, new research and little space

is given for negativity or critical discussion of existing research. Engage is the third most frequent function followed by Position. In the sciences in general, empirical research is deemed neutral and largely an extension of previous research. It is important to recall that, due to the limited number of papers analysed, these are only trends in the selected papers and cannot be generalised to all GEC papers.

Figures 5.14 and 5.15 are visual summaries of the proportions of citation functions in both time periods of constellation 5. In constellation 1, the use of citations to Support has increased in recent years, and also Engage, but to different extents. On the other hand, Signpost and Position have decreased over time. In constellation 5, all the function proportions have increased in recent years, except for the function Engage which dropped by more than 50% in the recent years to converge with constellation 1.

In terms of citation functions, Figure 5.11 shows that constellations 1 and 5 were quite different in the early years. Interestingly, the early period of constellation 5 diverged from the other three data sets. Form, stance and function analyses all reveal this time period of constellation 5 to be distinct. While Support remains the most dominant function, the remaining three functions occur in a different order. Notably, Engage is the second most frequent function in the selected papers from the early period of constellation 5. In fact, it is by far the most frequent function in one paper from constellation 5 (Beckerman and Pasek 1995).[2] Also noteworthy is that the same period contains more use of Position citations (16.97%, see Figure 5.11) in comparison to the two time periods of constellation 1. Unlike constellation 1, this constellation shows a diachronic increase in the use of the function Position, possibly because it deals with personal voices (constellation 5 is titled ‘Personal Voices’ in Thompson et al. 2017) and represents the social science-like component of the GEC corpus.

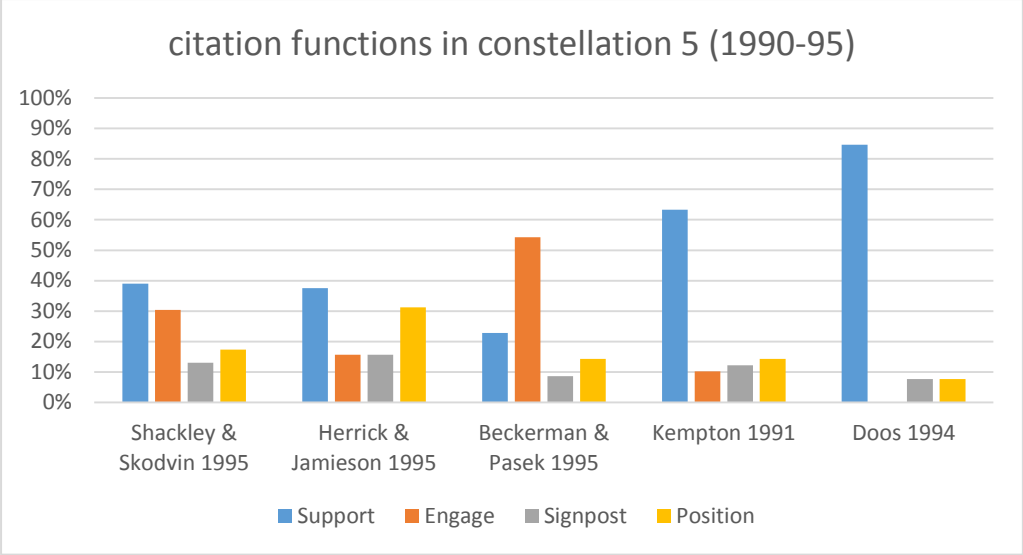

Figure. 5.14: Citation function ratios in the early papers of constellation 5

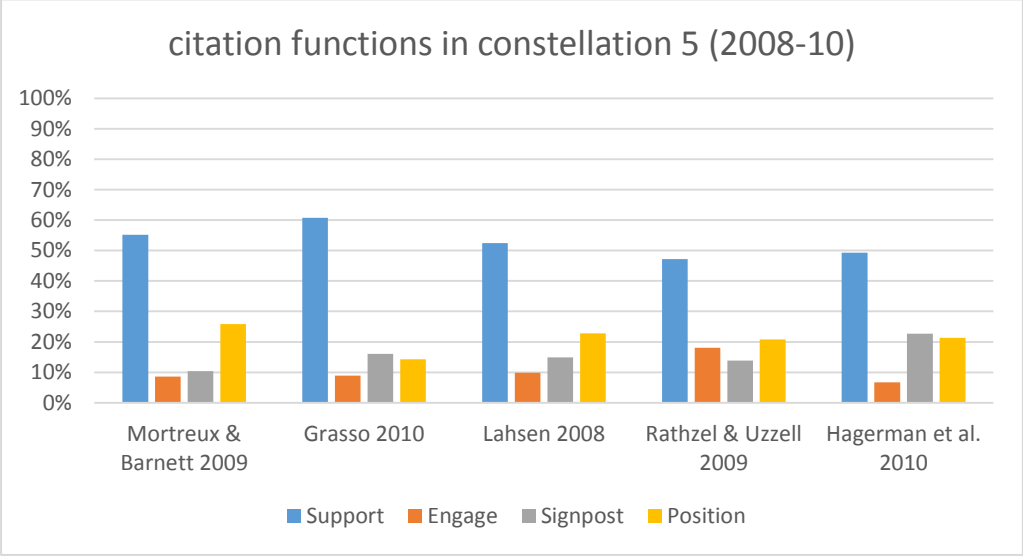

Figure. 5.15: Citation function ratios in the recent papers of constellation 5

Signpost as a function has diachronically increased in constellation 5, yet it remains less common than in constellation 1. Position has also increased in frequency in recent years. In general, Position is more common in constellation 5 than 1. There seems to be a convergence between constellations 5 and 1 towards a middle ground in terms of the functions Support and Engage. Figure 5.11 also shows divergence in terms of the functions Position and Signpost. Position and Signpost have decreased over time in constellation 1 while showing an increase in constellation 5. It seems reasonable that a function such as Position (to provide external viewpoints) would increase over time in constellation 5. This is because the nature of papers in this constellation predicts the

presentation of and engagement with the various voices which keep arising as the field grows and more views are introduced.

Constellation 5 is distinguished by its engagement with sources and is therefore expected to contain higher proportions of both Engage and Position and lower proportions of Signpost and Support compared to constellation 1. It must be observed that there is the diachronic decrease of Engage (critical dialogue with sources) in the selected papers from constellation 5. There are various possibilities for this decrease, some of which appear in form analysis and could be summarised as either 1) an editorial decision, 2) the inevitable consequence of the journal being more focused, or 3) a constellation style issue. However, the model needs to be tested on a larger corpus in order to verify on a large-scale any diachronic convergence or even divergence between the constellations.

Evident from constellation 5 is that one paper (namely: Beckerman and Pasek 1995) was found to be quite different to others in the same constellation, yielding different results, particularly regarding form and function categories. For example, it was the only paper in which Support was not the most common function. Instead, it was the only paper with Engage as the most common function. Furthermore, it was the only paper in the four sub-corpora that contained more integral than non-integral citations with approximately 60% of its citations being integral. The next chapter provides a textual analysis of this paper and compares it to another typical paper to show how and why it appears distinct.

To summarise, and as can be seen from Figure 5.11, the order of the citation functions in terms of the most-to-least frequently used citations is as follows for both time periods of constellation 1: Support, Signpost, Engage and Position, respectively. However, for constellation 5, the two time periods yield different results: in the early papers, the order of frequency is Support, Engage, Position and Signpost; in the recent papers, however, the order is Support, Position, Signpost and Engage. The relatively low frequency of the function Engage may suggest that the journal is less concerned with antagonistic approaches in the study of environment-related issues.

5.5 Revisiting stance and function: explicit stance, implicit stance and multi-functional citations

This section outlines synchronic and diachronic variation between the four sets of data in terms of citations with single stance (SS), citations with implicit stance (also referred to as 'multi-stanced' citations, hereafter MS) and multi-functional (MF) citations. The purpose of this section is to trace variation between the data sets and to contribute to what is known about the variation from the aforementioned analyses of citation form, stance and function. The findings from this section are linked to the characteristics of the two constellations (1 and 5) as reported in Thompson et al. (2017) and demonstrated above in sections 5.2, 5.3 and 5.4.

As a brief reminder, explicit stance or SS citations in this section refers to citations in which one explicit stance is identified (e.g. when a writer explicitly *contests* a source or an argument at the clause-level of the citing proposition as in Example 4.7). Implicit stance co-occurs with explicit stance in multi-stanced (MS) citations. Examples of MS citations are included in Section 4.2. As a reminder, examples include citations whereby writers explicitly *acknowledge* a source and at the same time implicitly *endorse* the source's contribution (see Chapter 4: Example 4.13 for a discussion) or when implicit *contest* is identified at the surface of explicit *acknowledge* or *distance* (see Examples 4.26 and 4.27, respectively). MF citations have been discussed extensively in the literature (e.g. in Harwood 2009 and Samraj 2013). They include citations with more than one function in the block (e.g. when a citation is included in Support of previous text but in Engagement with following text; see Example 6.10 in the next chapter).

The following figure (Figure 5.16) outlines the synchronic and diachronic variation between the four data sets in terms of SS, MS and MF citation proportions. The importance of this comparison stems from its relevance to explicit and implicit argumentation at the clause- (SS and MS) and block-levels (MF). I shall not go into detail about how MF citations relate to explicit or implicit argumentation at the block-level, due to time and space restrictions. However, it is sufficient to argue that writers draw links between propositions to show their overall stance (as noted by Thompson et al. 2017) or position on the cited arguments in the block. The writer's stance(s) or viewpoint(s) in the block is(are) taken as indication of how they employ the citation in the text (i.e. the citation role or function in the block). The block-level stance reflects

the citation function(s) and writers draw links, either explicitly or implicitly, between propositions in the block to reflect one or more functions. Thompson et al. (2017) note that writers sometimes guide their readers' interpretation through the text. Likewise, I would argue that writers can make the links between propositions explicit or implicit not only to (un)guide the readers' interpretation of the citation function(s) but also to forge the links between the functions either overtly or covertly.

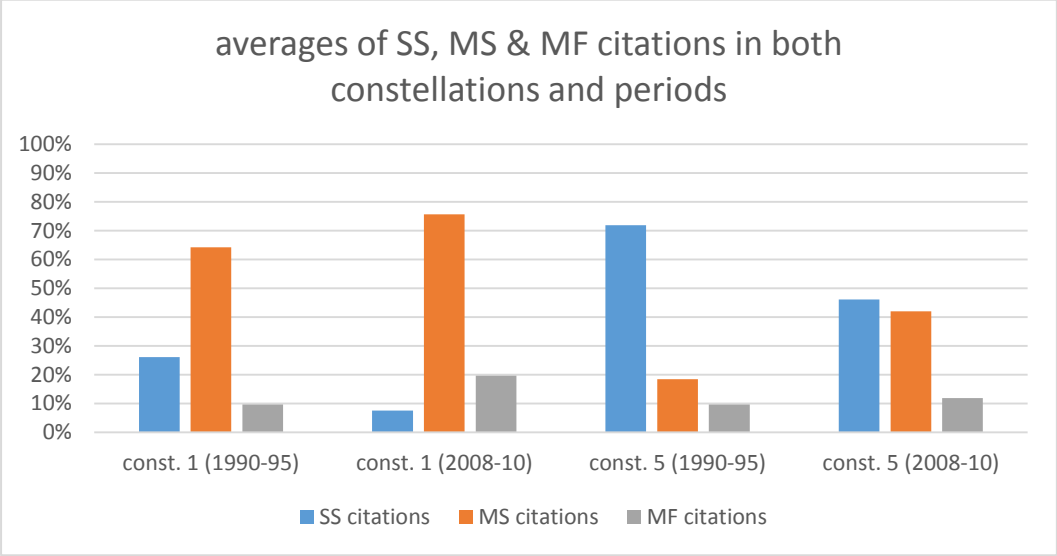

Figure 5.16: Averages of SS, MS and MF citations in the four data sets

Returning to the clause-level argumentation, the current thesis links explicit stance to explicit argumentation and implicit stance to implicit argumentation. As Figure 5.16 shows, in terms of stance (both SS and MS), the constellation 1 papers display high ratios of MS citations in both time periods whereas SS citations dominate in the constellation 5 papers. What is remarkable in both constellations is the diachronic increase of MS citations at the expense of SS citations. In other words, the early papers from constellation 5 were different from the other time periods in both constellations. As the time went by, constellation 5 started to converge with constellation 1 in terms of stance proportions (i.e. decrease in SS and increase in MS).

It is not surprising that constellation 1 papers include high ratios of MS citations, because this constellation is characterised by implicit argumentation. At the clause-level of citation, implicit stance is taken as an indication of implicit argumentation. What is striking, however, is that constellation 5 started to incorporate more MS citations in the recent years to converge with constellation 1. This finding is clearly related to the dominance, and diachronic increase, of non-integral citations in both

constellations. Implicit stance tends to occur in non-integral forms. Taken at face value, the writers in some non-integral citations might seem to simply *acknowledge* a source, but implicit stance analysis has shown that writers can, and often do, conceal their positive or negative stances towards the same source. Such occasions of implicit stance in non-integral forms contribute to the high proportions and diachronic increase of MS citations at the expense of SS citations in both constellations.

In the early years, the constellation 5 papers were different as they contained higher proportions of integral citations (see Figure 5.1) compared to the other constellation and time periods. The high proportions of integral citations in the early papers of constellation 5 matched a noticeable dominance of SS citations in the same period of the same constellation (compare the early period of constellation 5 in Figures 5.1. and 5.16). This confirms the finding that integral forms are associated with more apparent stance. Put differently, the expression of the writer's stance is more straightforward and easier to identify in integral citations (e.g. *Smith (2000) claims, fails to, demonstrates* or *compellingly argues*).

Moving on to citation functions, MF citations have increased to different extents in both constellations. In constellation 1, MF citations doubled in the recent years (from 9.67% to 19.66%). In constellation 5, MF citations also increased over time (from 9.66% to 11.91%) but less than in constellation 1. The increase in MF citations seems to suggest that the practice of citing became more complex in the recent years. In other words, writers in the recent years use citations to achieve a number of functions in the blocks more than their counterparts in the early years. The diachronic increase is slower in the selected papers from constellation 5. This could well be due to the nature of this constellation, as citations in this constellation seem to be associated with explicit argumentation and engagement with human perspectives (Thompson et al. 2017). Therefore, it seems reasonable to argue that the writers in this constellation prefer to directly engage, at the block-level, with the sources (e.g. to enter into critical dialogue with the source; see Chapter 6 for a case study of one paper from constellation 5).

The identification and analysis of MF citations emphasises the importance of the block as a new unit of analysing citations. Citation role(s) cannot always be identified at the clause-level of the citing statement. More context is often needed for an accurate identification of the citation function(s) with respect to the writer's argument, finding

and/or main idea (i.e. whether the citation Supports the writers argument in the block, Signposts the reader to an external source, shows a source's Position or is included to show the writer's Engagement with the source's views). This context (the block) is used to identify citation function(s) and also the links drawn by the writer between two or more external sources (see the following section).

5.6 The cited-cited relation

This section quantifies the cases in which the writers drew links between two or more external sources. Two links are identified here, as noted in Section 4.4; these are citation-support and citation-engagement. It is notable that these links are different from citation functions in the text as the latter accounts for the ways in which writers use cited material to perform different roles in the ongoing discussion (e.g. to Support the writer's argument). The cited-cited relation, while also identified at the block-level, include cases whereby writers either show similarity between cited studies (see Examples 4.28, and 6.3 in the next chapter), agreement between two or more sources (see Examples 4.38 and 4.43) or disagreement with previous research (see Example 4.43; i.e. how S3 relates to S1).

Overall, the writers draw 141 links between the 1186 citations. 82 of the links are drawn between the citations in constellation 1, whereas the remaining 59 links occur in constellation 5. This could possibly be due to constellation 1 papers attaching "value judgements" to climate change predictions and environment-related issues (Thompson et al. 2017: 19) and therefore making more links between cited sources. Scientific research tends to build on and extend previous empirical research, this extension may include drawing links between sources to show how the various studies compare or contrast. In the two constellations, 81 of the links show citation-support whereas the remaining 60 links show engagement between cited sources. See Table 5.1 below.

As Table 5.1 shows, both citation-support and citation-engagement have increased over time in the papers from the two constellations. The early papers contain 49 instances of both kinds while the recent papers contain 92 instances. This would suggest that recent papers are expected to draw more links between sources and align their research accordingly. Also, instances showing citation-support are more frequent than those showing citation-engagement in both constellations. One possible interpretation is that the selected papers tend to cite studies which are in accord with each other (i.e. when

citing, writers choose studies which are not putting themselves in conflict with their predecessors' research). One exception is the diachronic increase of citation-

Table 5.1: number of instances showing citation-support and citation-engagement in the four data sets

		Number of instances showing citation-support	Number of instances showing citation-engagement	Total number of instances in constellations 1 and 5
Constellation 1	1990-95	18	8	82
	2008-10	29	27	
Constellation 5	1990-95	12	11	59
	2008-10	22	14	

engagement in constellation 1 (from 8 to 27). This increase of citation-engagement suggests that the writers in constellation 1 (Quantification and Measuring) draw more negative links in the recent years (e.g. between studies of "specific sites of interaction between people and the environment"; as noted by Thompson et al. 2017: 19), showing points of variation or contrast between the cited studies.

5.7 Summary

This chapter has quantitatively investigated the occurrence of the 10 categories of form, stance and function in the selected papers. The synchronic and diachronic variations between the papers were documented in this chapter. The main findings can be summarised in the following points:

- Constellation 5 differs from constellation 1 in terms of all three aspects of analysis (form, stance and function) in the early papers.
- Form, stance and function analyses show divergence in the early years of the selected papers between the two constellations, and convergence in the recent years.
- The diachronic change was most noticeable in constellation 5. The early period of this constellation is different from the other data sets. However, the recent papers in this constellation show convergence with the constellation 1 papers.

- Figures 5.4, 5.9 and 5.14 show one paper from constellation 5 (Beckerman and Pasek's 1995) to be different from the other papers in the same constellation and from the constellation 1 papers. This is particularly the case with respect to form and function analyses. For example, it is the only paper to use more integral than non-integral citations (60%) and the only paper in which the function Engage occurs more often than Support. For qualitative examples, see Chapter 6.
- Papers in constellation 1 show higher proportions of multi-stanced (MS) citations. This finding can be justified when considering the observation of Thompson et al. (2017) that papers in this constellation score highly on implicit argumentation. Although the authors did not study citation practices and arrived at this observation by considering the ways in which writers draw links between propositions and expressing stance, this finding is consistent with their observation. Constellation 1 contains more instances of MS citations (e.g. Example 4.13 "*The impact … is unprecedented (REF)*") than constellation 5. By contrast, constellation 5 papers contained more SS citations, which can also be attributed to explicit argumentation. The fact that more SS citations were identified in the constellation 5 papers reflects a tendency towards explicit stance, which tends to be identified in integral citations. Figure 5.1 shows higher proportions of integral citations in the constellation 5 papers, compared to the other papers in constellation 1. The integral citations facilitate explicit expression of stance at the clause-level.
- In terms of the cited-cited relation, both citation-support and citation-engagement have increased over time in the selected papers from both constellations. It is natural for recent papers to draw more links between previous sources and to align their research accordingly. The analysis of constellation 1 papers reveals a greater diachronic increase in such links than does the constellation 5 papers. In fact, Thompson et al. (2017: 19) observe that papers in constellation 1 "attach value judgements" to climate change predictions and related issues. In terms of citation practices, it seems that writers in this constellation tend to highlight how previous studies relate to research of their predecessors more than those writing in constellation 5, who seem to prefer to engage directly with sources.

Notes

1- The average ratio of each feature (e.g. integral form) for each sub-corpus was derived by adding the ratios of all the papers in that sub-corpus, divided by the number of papers in that sub-corpus.

2- Due to the limited number of papers included in the analysis, it is to be expected that the outlying results of one paper would skew the overall results of a particular sub-corpus. This is because each paper (Beckerman and Pasek 1995 as an example) represents one-fifth of the papers in the early period of constellation 5. However, as noted above, the aim of this thesis and this chapter in particular is not to compare different files (as in Thompson 2001). Therefore, the figures have not been aggregated to study citation density, or the distribution of stance and function categories. Instead, the main aim is to show proportions of categories of stance and function which together form the conventional and unconventional combinations, and which can be realised in both integral and/or non-integral citations. To generalise the quantitative results and for them to be representative of the constellations (or GEC in general), a more representative sample is required. At present, it suffices to recall that the main aim of this thesis is to propose a model, with which to study the role of citations in the citing text. Another aim is to make the proposed model replicable, even if the quantitative results obtained from applying the proposed model to the selected papers may not be generalised or indeed representative of all GEC papers.

Chapter Six

Textual Analysis of Citation Blocks: A Comparative Study

6.1 Introduction

The previous chapter provides evidence of variation between the two constellations and time periods. The quantitative comparisons showcase variation in terms of the occurrence and spread of the form, stance and function categories in the selected papers from the two constellations. In preparation for this chapter, my own intention was to select from the two constellations and time periods two papers that were as different as possible, as reflected in the quantitative comparisons conducted in the previous chapter. The impetus for this decision is to demonstrate the rhetorical effects of the analyses I had conducted, relating particularly to the combination of the form, stance and function categories.

In order to achieve this goal and reflect on citations from the selected papers, the current chapter applies the proposed model of analysing citations found in two papers derived from the two constellations and the two time periods (viz. Beckerman and Pasek 1995 and McAlpine et al. 2009; see the list of papers analysed in this thesis at the end of the reference list). In applying the proposed model, the current chapter uses close-reading as a method for investigating and comparing the citation practices found in the two papers.

As noted in Chapter 5, the quantitative analyses showcase one paper from constellation 5 (namely: Beckerman and Pasek 1995) to be markedly distinct from the other 19 papers. For example, this paper differs from the other papers in terms of the results of its form and function analyses. Integral citations account for 60% of all the citations in this paper whereas non-integral citations are dominant in all the other papers (see the first five figures in Chapter 5). The function analysis displays the greatest difference, as this is the only paper where Engage is the prevailing function (accounting for 54.28% of citation functions). On the other hand, Support is the dominant function in the other 19 papers. For this reason, this specific paper was the obvious choice for the constellation 5 paper.

In search for a contrasting paper from the other constellation and time period, the figures in Chapter 5 showcase McAlpine et al.'s (2009) paper to be noticeably different from Beckerman and Pasek's paper (compare Figures 5.3, 5.8, 5.13 with Figures 5.4, 5.9 and 5.14, respectively). In short, these two papers were selected for analysis as they not only demonstrated the greatest variation in the occurrence and spread of the 10 analytical categories but also represented two different constellations and time periods. The figures from the previous chapter show dominance of the *acknowledge*-to-Engage combination in the paper by Beckerman and Pasek. On the other hand, the same figures demonstrate a preference for the *acknowledge*-to-Support combination in the paper by McAlpine and his colleagues.

At the end of this chapter, I will discuss the implications of the analysis for GEC as an interdisciplinary journal and the IDRD project (as reported in Thompson et al. 2017). In summation, both these papers provide quantitative and qualitative evidence of variation between the two constellations and time periods (see Sections 6.2, 6.3 and 6.4). The current chapter uses the block as the main unit of analysis to demonstrate variation in terms of citation practices between the two papers. The analysis stresses the importance of combining the form, stance and function categories at the clause- and block-levels to demonstrate the variation between the two papers.

This chapter is divided into four sections. The next section (6.2) introduces McAlpine et al.'s paper, as a typical paper from constellation 1. It also introduces the main sequences identified in citation blocks in that paper and the manner in which the authors of that paper cite external research. The following section (6.3) introduces the distinct paper (Beckerman and Pasek) with a detailed discussion of most of the citation blocks identified in that paper. This section starts with a brief overview of the content and purpose of that paper. Then, the simple sequences of citations are introduced with some examples. The same section also includes examples which demonstrate the complexity of dialogue within citation blocks. The last section (6.4) summarises the main arguments and compares the findings from the two papers. This section also discusses the findings in relation to GEC as an interdisciplinary journal and the IDRD project.

6.2 McAlpine et al. (2009)

As noted in Thompson et al. (2017) and in Chapter 3, papers in constellation 1 represent the scientific component of the GEC corpus. McAlpine et al.'s paper is a typical

example of the constellation 1 papers. The figures in Chapter 5 display variations between the selected papers from the two constellations and time periods. However, the same figures show this paper to be similar to some extent to the other papers from both constellations. For example, the vast majority of citations in both constellations are non-integral and feature Support as the most common function. In terms of stance analysis, the figures also show a dominance of *acknowledge*, followed immediately by *endorse* (except in two papers from the early period of constellation 5; viz. Beckerman and Pasek 1995 and Kempton 1991).

The purpose of this section is to analyse McAlpine et al.'s paper (as a typical paper from constellation 1) in order to highlight the main stance and function combinations and the sequences identified in citation blocks. McAlpine et al.'s (2009) paper contains 154 citations. The citation blocks in McAlpine et al.'s paper tend to be short, mostly containing two statements (see Examples 6.1 and 6.2). However, the blocks on some occasions contain three or four statements, but in general they are identified in the same paragraph in which the citations occur. Due to time and space restrictions, it is difficult to analyse all or most of the citations. Therefore, I have selected examples which demonstrate different sequences of moves that I have identified in the citation blocks.

The authors of this paper discuss the policies associated with beef production and consumption and call for action against the increasing demand for beef based on evidence collected mainly from Queensland (Australia), Columbia and Brazil. The authors also discuss the impact of increased beef consumption from various perspectives, including that of the environment, based primarily on evidence collected mainly from Queensland, Australia. They perceive the production of beef as a serious problem with various repercussions negatively impacting the environment. Evidence indicates (as they suggest in the abstract) that increased consumption of beef is causing "regional and global change, and warrants greater policy attention" (McAlpine et al. 2009: 21). The authors propose four policy imperatives to mitigate such impact on the environment. See the following table (6.1) for a summary of the different sections and a brief description of the content of each section.

The previous chapter reveals a preference in this paper for using non-integral citations (representing 92.81% of the citations). Also, 52.14% of the citations show *acknowledgement,* followed by 46.07% showing *Endorsement* and 1.78% showing

distance, while none of the citations show *contest* in this paper. The reason behind the high ratios of *endorse* immediately after *acknowledge* is the identification of implicit *endorsement* in this paper. In most examples, such as Example 6.1 (which is quoted in Chapter 4 as Example 4.13), explicit *acknowledge* is identified in addition to implicit *endorse*. This contributes to the high ratios of *endorse*, as a stance type.

Table 6.1: Section labels and summary in McAlpine et al. (2009)

Section number	Section label	Summary of section content
1	Introduction	The writers introduce the environmental effects of the increasing consumption of meet. This section contains Citations 1-31.
2	Theoretical framework	The writers highlight key areas of intervention "to mitigate and ameliorate the negative impacts of beef production on the regional and global environment" (p. 23). This section contains Citations 32-33.
3	Queensland case study	This section is further divided into five sub-sections on the history and drivers of change in this area, land-use and cover changes, regional impacts, global impacts and policy responses. This section contains Citation 34-52.
4	Colombian case study	This section is also divided into five sub-sections similar to those in the previous section but in relation to Columbia. This section contains Citations 53-80.
5	Brazilian case study	This section is also divided into five sub-sections similar to those in the previous two sections but in relation to Brazil. This section contains Citations 81-115.
6	Synthesis of case studies: environmental, human and policy dimensions	In this section, the authors summarise the previous case studies and introduce four imperatives to help mitigate the environmental effects of beef production. This section contains Citations 116-154.
7	Conclusion	The writers summarise the main arguments. This section contains no citations.

Support is the dominant function (75.6%). There are some possible explanations for the high ratios of *endorsement* and Support. The first explanation is that the authors of this paper tend to focus on studies that Support their own arguments in the same blocks (see Examples 6.1 and 6.2) to persuade the reader of the need to tackle the issue of increasing consumption of beef. Almost all the papers they cite have found similar results, whether it is in Australia or elsewhere. Another explanation is that the writers frequently use self-citations where the same authors, or some of them, have conducted similar research in the same or in other areas (e.g. see the second citation of Example 6.4).

In Chapter 4 (Section 4.2), I argued that a citation may be implicitly *endorsed* in a number of ways. For example, if the cited message is presented as factual, true, trustworthy and/or derived from a trusted source, then it is implicitly *endorsed*. This paper contains many examples of implicit *endorsement* of the cited propositions. To exemplify this argument, consider the following examples in their short blocks. The first citation in this paper (Example 6.1) occurs in a two-statement block. What is immediately obvious is the lack of explicit signals of argumentation at the block-level (e.g. *However, Also, Indeed, Moreover*). In Example 6.1, the writers do not explain or make explicit, the steps involved in argumentation or the links between phrases and propositions (see Thompson et al. 2017 for examples of explicit and implicit indices of argumentation and the writers' stance).

Example 6.1:

> [1] *The **impact** of humans on the global environment **is unprecedented** (REF).* [2] Global climate change, deforestation, dwindling water resources, desertification and loss of biodiversity are all **symptoms of human-accelerated environmental change**.

The writers in this example explicitly *acknowledge* the source of information and implicitly *endorse* the cited argument, as it is presented as true and from reliable sources such as IPCC (Intergovernmental Panel on Climate Change). In other words, the writers in S1 are presenting this piece of information as a universally-accepted fact, (i.e. something that is known to be occurring). See the analysis of Example 4.13 for further details on how implicit stance is identified. S1 is cited in Support of S2. The writers in S2 aver that the *symptoms* (which provide anaphoric reference to *impact* in S1) are examples of human-induced environmental impact. In other words, S1 is cited in Support of the writers' argument in S2.

This citation exemplifies the first sequence of citation in this paper: 'cited fact + Averred elaboration'. In this pattern, the writers tend to *acknowledge* (and implicitly *endorse*) a statement of fact in the first statement in the block. The second statement, which is averred, elaborates on the cited fact in the first statement. The writers' argument in the second statement is in accord with the sources' statement in S1. Therefore, S1 is cited in Support of the writers' argument, which is deferred to the second statement. The Support is usually implicit. The writers often provide anaphoric reference in the second statement to an entity mentioned in the first, or use certain lexical items in the second statement to refer to previously cited lexis (compare *symptoms* and *impact* in Example 6.1).

The second sequence identified in this paper also contains two statements: 'Averred fact + Cited elaboration'. Consider the two statements in Example 6.2. The cited statement in this example (S2) includes a self-citation, whereby at least one of the authors has reached similar findings in their previous research.

Example 6.2:

> [1] Cattle have also played an important role in transforming Colombian landscapes since their introduction in the early 1500s (Fig. 4). [2] ***This long-lasting impact*** *on Colombian ecosystems* ***continues to date****, with the expansion of cattle grazing in frontier regions such as the Amazon and Magdalena* (Etter et al., 2008).

In S1, the writers aver a statement of fact: that *cattle have also played an important role in transforming Colombian landscapes* and that this impact dates back to the early 1500s. In the self-citation (S2), the cited information is presented as true, something which started in the past and *continues to date*. In S2, the use of *this* and *long-lasting* provide anaphoric reference to the prolonged cattle impact in S1 and in a way verifies the cited period in time referred to in S1. There is no overt indication of clause relations in Example 6.2 (e.g. *Moreover, In fact, Because, But, Nevertheless*).

The writers in S2 not only *acknowledge* the source of the information, they also *endorse* it implicitly since they evaluate it as true, or as a fact. At the clause-level of S2, there is no explicit marker of *endorsement* (e.g. *as noted by* or as *Etter et al. demonstrate*). Therefore, implicit *endorsement* is identified together with explicit *acknowledgement* in such examples to show how writers may conceal their stance towards the research they cite behind neutral words. *Acknowledgement* in this sense may be treated as contractive. This citation is used in Support of the averred information in S1.

The blocks in this paper are usually limited to one or two statements before or after the citing statement as in the previous two examples. This could be attributed to the high number of citations in this paper (154 citations). The writers cite many facts, thus there is less space to be devoted to explaining or commenting on each citation. In addition, the fact that the authors of this paper tend to cite statements of fact means that it is unlikely that the authors will comment on or Engage with those facts, whereas those citing arguments are more likely to engage with the cited propositions.

Some blocks in this paper contain more than two statements. These blocks contain sequences that are not as common as the first two. For instance, the next block (see Example 6.3) contains four statements. This example contains two citations in one block. Both citations (in S2 and S3) are cited in Support of S1. The writers cite some programmes which have promoted beef consumption. The cited programmes Support (i.e. provide evidence from the literature for) the information in S1 although the writers do not support those programmes or the government agendas which promote them, as can be seen in S4. The two citations act in two different ways. They Support the first statement and at the same time show Engagement with the writers' argument in following text (S4). The writers negatively evaluate such programmes as being *not environmentally responsible*.

Example 6.3:

> … [1] To mitigate this decline in per-capita consumption, **government agencies and the beef industry have actively promoted beef products as healthy and safe**. [2] ***The Australian government, for example****, has a long history of trade missions* ***promoting the consumption of Australian red meat in Asia and the Middle East*** *(Australian Bureau of Agricultural and Resource Economics, 2007).* [3] ***Similar programs exist in the United States****, such as the Beef Promotion and Research Act of 1985 (USDA, 1985).* [4] **We argue that such industry and government agendas are not environmentally responsible** as 'saving and repairing the planet' becomes the new global imperative.

This block exemplifies the third sequence identified in this paper: 'Averred fact + Cited examples + Criticism'. In this sequence, the writers aver a statement [of fact] in the first statement. In the second statement (and also the third in Example 6.3), the writers cite examples elaborating on the averred fact in the first statement. In the last statement, the writers criticise the cited examples. S4 shows the writers Engagement with the actions, not propositions, taken by the Australian and American governments as cited in S2 and S3. This sequence is uncommon in this paper as most of the citations in this paper are

cited in Support of the writers' arguments in the block, with no Engagement with the cited propositions.

At the clause-level of S2, the citation displays *acknowledgement* of the source of the information about the Australian government actually promoting its red meat's consumption (including beef). The same citation also implies the writers' *endorsement* of the cited information in two ways. First, this piece of information is presented as true. The presentation of the information in S2 as true does not mean that the writers agree with the cited actions for promoting red meat's consumption; rather, it means that the information is presented as an existing fact (a bald fact or an issue that needs to be addressed). Second, the information is obtained from a reliable source (Australian Bureau of Agricultural and Resource Economics) which boosts the information's credibility. Therefore, it is more likely to be accepted by the readers (as a given fact, at the immediate clause-level), even though the writers do not wish the reader to support the source's actions (see S4). Put differently, by 'accepted' I do not mean that the reader should, or should not, agree with the cited policies. I mean rather that the reader should find the source reliable enough to produce such information, as a given fact. The writers then position the reader to be critical of the cited policies or actions in S4. This is one of the examples that show the dialogic power of the Engagement system (Martin and White 2005). What matters here is not the clause-level analysis of stance (i.e. the writer-author evaluation); rather, it is the function and role of citations and how the block-level evaluation positions the reader to be critical of such programmes (i.e. the writer-reader interaction).

It should be noted that the second citation (in S3) performs a similar function to the first citation (in S2); both being cited in Support of the writers' argument in S1 and at the same time as being cited in Engagement with the writers' view in S4. The citation in S3 provides an additional example of beef promotion programmes. However, it might be argued that the second citation plays an additional role which is to Signpost the reader to a similar program that exists in a different context (in the USA, not in Australia, Columbia or Brazil). This block demonstrates citation-support between S3 and S2 (*Similar programmes*).

Another sequence is illustrated by Example 6.4 below. The first citation in this example is conceded and the second is asserted. The writers evaluate the first citation as valid

and the second as more valid (Thompson and Zhou 2000). There are no explicit markers of *endorsement* (e.g. evaluative lexis such as *compellingly, successful, demonstrate*) at the clause-level of the first or second citations. In the second citation, which is a self-citation, the writers reproduce a piece of information that two of them had found in previous research. The third citation (in S2) is a clear example in which the writers *distance* themselves from the cited proposition. In other words, the writers' choice of the verb *claims* suggests that the writers accept no responsibility for the cited claim, thus maximising the social distance between the writers and authors and also allowing for the formation of alternative reader views. At the block-level however, S2 supports the counter-assertion in S1, though that support is implicit. This block shows that the writers are less concerned in making their argumentation explicit. For example, there is no explicit connection (e.g. cohesive device) which links S2 to S1 or S3 to S2. There is evidence in the literature (Thompson et al. 2017) that constellation 1 papers score highly on features associated with implicit argumentation. McAlpine et al.'s paper is a typical paper from constellation 1 where the writers display a lack of concern to make the links between the presented propositions, explicit.

Example 6.4:

> [1] ***While there is uncertainty in estimates*** *of carbon stocks in tropical forests (IPCC, 2007b),* ***the effects*** *of tropical deforestation on greenhouse emissions* ***are substantial*** *(Fearnside and Laurance, 2004*). [2] ***Subak (1999) claims*** *feedlots, with their intensive, high technology management, produce nearly twice the CO_2 equivalent per animal as compared to cattle grazing in pastures.* [3] The increasing use of grain and soy meal for feedlots **adds to anthropogenic greenhouse gas emissions**, with nitrous oxide emissions from agricultural soils contributing a further 6% to total emissions.

The writers present the cited information in the second clause of S1 as a fact and so it is asserted (*the effects … are substantial*). S2 exemplifies the *distance*-to-Support combination. The source's (Subak) claim about the high production of CO_2 per animal in feedlots is consistent with the argument that *the effects … are substantial* in S1. It also Supports the writers' argument in S3 that the increase in grain and soy usage in feedlots adds to *anthropogenic greenhouse gas emissions*. Even though S1 starts with an explicit index (*While*) which shows how the first clause relates to the second (in a concession and counter-assertion sequence), the writers make no explicit links between the three propositions at the block-level.

The citations in this example belong to the last sequence: 'Concession and counter accretion + Cited argument + Averred fact'. It is interesting that the citations (including

the concession, counter-assertion and the cited claim) are cited in Support of the averred statement in S3. When writers concede or *distance* an argument (as in the first clause in S1, or in S2), it might be expected that they will Engage with these conceded or *distanced* arguments later in the text. However, the authors of this paper seem to cite research that supports their arguments in the citation blocks, even if they are uncertain about the cited quantity as the clause-level analysis of S2 indicates (that *feedlots … produce nearly twice the CO_2 …)*.

In short, the authors of this paper use citations which Support their arguments at the level of the block. The citations draw heavily on universally-accepted facts. The analysis also shows the blocks to be relatively short in this paper. Even though the writers do not make explicit links between propositions, their position towards the research they cite and the citation function can easily be identified. This could be partly attributed to the short length of the blocks and tendency to cite facts, as opposed to arguments. We now turn to the more complex issue of the constellation 5 paper, where there are fewer citations but more explicit links between the clauses, propositions and blocks.

6.3 Beckerman and Pasek (1995)

The authors of this paper tackle the issue of the international allocation of tradable carbon emission permits. They are very critical of the existing theories of justice relating to the distribution of Tradable Permits (hereafter TPs) and their application. According to the authors, the existing theories are unhelpful in addressing the problem of global equitable distribution for a number of reasons. Particularly, the ethical motives for the existing theories apply to individuals rather than nations. For example, they criticise theories of justice that are based on equal per capita emission. Such theories call for the allocation of the total number of TPs between nations on a per capita consumption-and-emission basis. The authors believe that the existing theories are limited in scope and apply to individuals or on national levels, but not to TP distribution between sovereign countries.

In response to the perceived limitations of the existing theories, and from a purely political economist perspective, the authors argue that all countries should come to an agreement on the distribution of tradable CO_2 emission permits as long as this agreement ensures justice and is satisfactory for all participating countries. To achieve

this goal, the authors call for bargaining and negotiations over the allocation and distribution of permits and the replacement of current methods with new alternatives such as the involvement of international agencies to sell or lease the tradable permits.

The authors discuss and critique the various approaches and rules for allocating TPs throughout the different sections of the paper. Table 6.2 highlights the different sections in Beckerman and Pasek's paper, with a brief summary of each section.

Table 6.2: Section labels with a brief content summary in Beckerman and Pasek (1995)

Section number	Section label	Summary of section content
1	Introduction (not specifically labelled)	The writers critique the concept of equity (contains the first two citations)
2	The contribution of theories of justice	The writers critique the existing theories of justice. The proponents of the theories are divided into four groups based on the perceptions of their originators. The proponents of the first group are communitarians who do not believe in international distributive justice. The proponents of the second group are utilitarians who call for international redistribution based on world usage and need for the permits. The third group are the contractarians, including Rawls (1972), who are mainly concerned with international environmental protection. The fourth and final group, including Barry (1973) and Beitz (1988), critique some aspects of contractarianism and mainly Rawls' formulation and call for modified approach of Rawls' theory. The above theories are mainly criticised by the writers for irrelevance and inapplicability to real-world problems and their inability to redress and bring about fairer distribution of tradable permits between sovereign nations. This section contains Citations 3-18.
		This section critiques four equitable rules for permit allocation and contains Citations 19-27. The title of this section prospects (Sinclair 1991; Rego 2001) the writers' disapproval of the cited rules. The

3	A critique of 'equitable' rules that have been considered	writers heavily critique the rules beyond the clause-level and particularly at the block-level. The block can be the paragraph in which the citation(s) appear(s) as in Examples 6.5 and 6.6. It can also include two paragraphs, as in Examples 6.7 and 6.8, or a number of paragraphs preceding and succeeding the citing statements as in Example 6.9. This section contains Citation 19-27.
4	The basis for a quasi-equitable bargain	The writers outline their approach to distributing tradable permits. This section contains only one citation. Based on the writers' criticism of the previous theories and rules of international justice, they devise their own principles in this section and use the last citation in two contradictory ways (see Example 6.10).
5	Conclusions	Summary of the main arguments (contains no citations)

Overall, this paper contains only 28 citations, occurring in 12 blocks. Two of the blocks exemplify the function Support and are in one way or another similar to citations in the other papers. Therefore, they will not be examined in this section. The remaining 10 blocks include citations which show the extent to which the authors of this paper are critical of the research they cite. However, four blocks show the overlap between blocks and are in a sense similar to each other. Therefore, two blocks will be analysed (see Examples 6.11 and 6.12). In total, eight blocks will be examined in this section for the co-occurrence of the form, stance and function categories in order to demonstrate the textual effect of applying the proposed model and how it can reflect variation between different texts.

The remainder of this section is divided into three parts. The first part introduces and exemplifies the simple sequences of citation that have been identified in this paper, noting the extent to which they differ from the sequences identified in the previous paper. The following part introduces a more complex issue relating to the complexity of dialogue between the interlocutors (the writers, the cited authors and the reader). This part of the section includes three blocks. The complexity of dialogue in these examples will be revisited towards the end of this section. The last part of this section highlights another aspect relating to the complexity of citation in this paper: that is, the

overlap between blocks. Examples in this part of the section are viewed chronologically because the blocks they represent are not totally discrete from one another. In other words, additional information about some blocks is obtained by looking at their preceding blocks (see Examples 6.11 and 6.12).

As noted above, the figures in Chapter 5 showcase Beckerman and Pasek's paper to be markedly distinct from the other 19 papers. Most of the citations in this paper belong to the *acknowledge*-to-Engage combination. In general, the authors of this paper cite arguments to contradict them in the blocks. It will be noted that unlike McAlpine et al., the contradiction is explicitly signalled, mostly by connecting words (e.g. *But, However, for Example*) or lexical repetition. The examination will also make reference to the role of phraseology in making explicit the links drawn between propositions, in the more complex examples. The discussion of the examples is given at the end of this section and in the next section. In the next section, these examples are taken into account, as are the citation practices in the two papers.

The first identified sequence is exemplified in the following two examples (6.5 and 6.6): 'Cited research + Criticism'. In this sequence, the writers cite arguments (as opposed to 'facts' in the previous paper) in the first statement. The second statement in this sequence is averred and introduces the writers' disagreement of the arguments they cite. In Example 6.5 below, the writers cite three sources in the first statement. This interrogative statement is divided into two clauses. The writers criticise the three sources in the second statement (the citing statements are italicised in all the examples and the linguistic markers of topic, attribution, stance and function are given in bold and then discussed in the description of the examples).

Example 6.5:

> [1] *What about contractarian theories of justice, such as their latest formulation by John* ***Rawls****,*[fn] *which* ***is frequently cited as providing some guidance*** *in the international allocation of the burden of global environmental protection?*[fn] [2] The short **answer** is that **Rawlsian theory has little bearing on the problem in hand**.

At the clause-level or even at the interrogative-statement-level, stance analysis is insufficient to reflect the role of both citations. In S1, the writers do not openly pass judgement on either the first or second citation. In the second citation (also in S1), the writers *acknowledge* proponents of the first sources' view. The second citation

Signposts readers to some studies which cite and endorse the first source (*frequently cited as providing some guidance in …*). In S2, however, the writers criticise the first source, and *contractarian theories of justice*. They present their criticism in the form of an answer to the question they pose in S1, describing such theories as irrelevant. The positive evaluation of Rawls' theory by the other sources in the second part of S1 is also irrelevant to the writers. This is because it concerns global environmental protection while the writers discuss international justice relating to TP distribution. In other words, if Rawls is wrong then the people who cite him positively are also wrong. Both citations in Example 6.2 belong to the *acknowledge*-to-Engage pattern.

Exemplifying the same sequence is Example 6.6 below. Taken in isolation of its block, the citation in S1 seems to Signpost readers to Barrett's list of the criteria. However, the subsequent statement (S2) clearly reveals a different function of this citation. The writers in S2 are involved in critical dialogue with Barrett's criteria in S1. The first indication of criticism is *But* in S2. The pronouns *them* and *they* in S2 both refer to the criteria in S1. The writers explicitly argue that such criteria are *ad hoc* and unable to stand up to scrutiny.

Example 6.6:

> [1] *The main criteria that have been proposed, on 'equity' (and on other) grounds, for allocating TPs between countries* ***have been enumerated by Barrett.***[fn] [2] **But since none of them faces up to the fact that** mainline theories of distributive justice hardly apply to the question of distribution between national governments **they either misrepresent these theories or they rely on rather ad hoc principles which do not stand up well to scrutiny**.

Example 6.7 below exemplifies a new sequence in which, unlike the previous sequence, the block starts with averred statement(s). The sequence in this block contains 'Averred statement + Cited elaboration + Criticism'. The cited arguments in such a sequence are inserted as examples of the averred proposition in the first statement. The writers conclude this sequence with an averred disagreement to the cited arguments. This block consists of two paragraphs and contains three citations occurring in S3, S4 and S5. All the citations are presented in support of each other. That by no means suggests that they also Support the writers' argument at the block-level.

Example 6.7:

> [1] One **very commonly suggested rule** is to allocate the total number of permits amongst countries on a per capita basis. [2] In fact it was explicitly proposed in an early draft of the (International) Framework Convention a revised version of which was agreed at the 1992 Rio 'Earth Summit'. [3] ***According to this convention*** *emissions were to converge to a common per capita level over the course of the years.*[fn] [4] *One of the experts on the international negotiation of conventions to reduce carbon emissions,* ***Grubb, says of this rule that*** *'The moral principle is simple, namely that every human being has an equal right to use the atmospheric resource'.* [fn] [5] ***Other commentators, too, have insisted on*** *the ethical force of this rule.*[fn]
>
> [6] **However**, in the light of the above discussion **this principle appears to be far less compelling than it might appear at first sight**. …

S1 is averred by the writers even though it includes an attributed clause (*to allocate the total number …*). This statement includes no citation. In S2, the writers *acknowledge* another source adopting this rule, but with no citation. S3, S4 and S5 contain the first, second and third citations. The writers build up the opposition argument in the first paragraph of this block (S1-S5). At the beginning of the second paragraph, S6 signals a reversal in the writers' argument. The signs of this contrast are *However* and *less compelling*. In practice, S6 shows the writers' Engagement with the cited views.

As for the writers' stance in the first citation in S3, the explicit stance is *distancing* on the part of the writers. The use of *According to this convention* suggests that what follows is attributed to this convention and the source in which it appears, and that the writers share no responsibility for its reliability. The next two citations (in S4 and S5) provide citation-support for the first citation (in S3). The sources in S4 and S5 are both *acknowledged* by the writers. The use of the neutral verb *says* in S4 indicates no evaluation of the source (Grubb), even though the writers present this source as an *expert* in the same statement.[1] S5 simply *acknowledges* other proponents of the "Equal per capita allocations" rule.

The writers withdraw from criticising these sources in the citing statements. They choose to Engage, in critical dialogue, with these sources in the block (see S6). In fact, they Engage with the principle including the above three citations in the following four paragraphs. However, the block only includes S6 because it includes the first signs of the citations' function. The citation in S5 shows citation-support to S4 and S3. S4 is also another example of citation-support to S3. The cited-citing (function) and cited-

cited relations can be summarised visually in the following figure (6.1) which contains S3-S6.

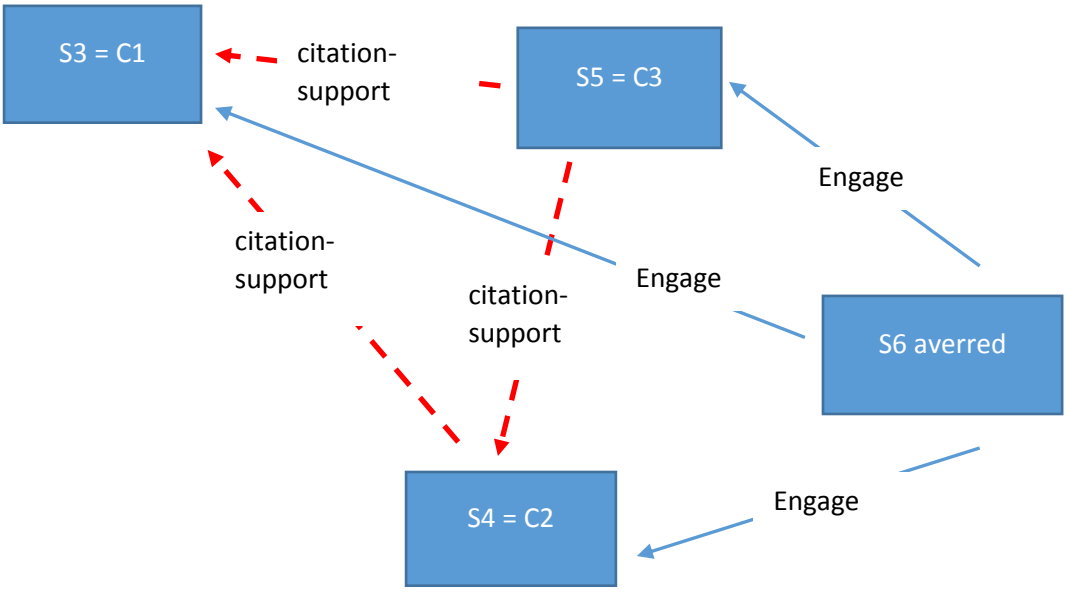

Figure 6.1: Visual representation of the cited-citing relation (citation function) and the cited-cited relations at the block-level of Example 6.7.

Unlike the previous example (6.7) whereby the cited arguments exemplify and elaborate on the writers' argument in S1, the citations in Example 6.8 are included to Engage with the writers' averred argument in the first statement. This example exemplifies a more complex sequence which demonstrates the complexity of dialogue between the interlocutors (i.e. the writers, the cited authors and the reader). This sequence contains 'Averred argument + Cited opposition + Criticism'. Overall, it could be argued that this example contains elements of the writer-author evaluation ('Status', Hunston 2000, 2011) and the writer-reader interaction ('Engagement' as within Martin and White's 2005 Appraisal theory).

Example 6.8:

> [1] **Negotiations** about the initial distribution of permits will**, of course,** be dominated by the **self-interest** of participating nations.[2] ***Nevertheless***, ***it is often maintained*** *that considerations of international* ***equity*** *will also play some part, even if only as providing 'focal variables', to use Barrett's term, around which the negotiations can be centred.*[fn]
> [3] ***Indeed***, *explicit* ***recognition*** *of the role of* ***'equity'*** *considerations* ***is contained in*** *the Framework Convention on Climate Change agreed at the 1992 Rio de Janeiro World Environmental Conference (INC 1992).*[fn]
>
> [4] **However**, the widely **conflicting interests** of different countries as regards the extent and form of measures to restrict carbon emissions **makes it unlikely that genuine equity**

> **considerations** will play **any role** at all as compared with naked bargaining, as the March/April 1995 Berlin conference on the subject has demonstrated. [5] **Furthermore**, as will be argued below, **it is highly unlikely anyway that** any widely accepted ethical principles could be found that have any bearing on the equitable international allocation of tradable emission **permits**, in spite of the impression to the contrary given by discussion of various 'equitable' allocations of emission permits. [6] The main reason for this is that **the discussion implicitly assumes that** there are some clear principles of distributive justice between nations that are a simple extension of those relating to distributive justice between individuals within any nation. [7] **But**, as is argued below, **this assumption is quite unwarranted**.

This block comprises two citations (in S2 and S3).[2] The block starts from sentence 1 (S1) and ends on and includes sentence 4 (S4). Sentences 5-7 are only included to show how the writers elaborate on their criticism of the sources' views. In other words, the citations' function in this example can be identified from the first four statements and S5-S7 are provided as additional evidence of the writers' Engagement with the citations in S2 and S3.

In S2 and S3, the writers cite studies which oppose the averred claim in S1, about the importance of the self-interest of participating nations in the negotiations over the distribution of TPs. The use of *of course* in the first averred statement (S1) sets up a concessive relation between S1 and S2. S2 and S3 are both attributed. S3 acts as evidence of the attributed claim in S2 and thus shows citation-support to S2. In S4, the writers oppose and respond to the information in S2 and S3 whilst confirming the information given in S1. The use of adversative conjuncts (or conjunctive adjuncts as in Halliday 1994) such as *However* in S4, clearly indicates a reversal in the direction of argumentation; not only to Engage with the attributed views in S2 and S3 but also to confirm the averred argument in S1.

Example 6.8 also demonstrates the need to stress the role of phraseology in citations and how it is picked up in the restatement of the writers' argument in S4. As can be seen from Table 6.3 below, the lexical repetition (e.g. in the restatement of *interests, equity and role*) acts as a cohesive chain that links information in the averred statements with information in the attributed statements in S2 and S3 as well as the averred concession in S1. Highlighted in yellow, green and blue are lexical items which are interrelated. For example, the *conflicting interests* (highlighted in yellow) in S4 picks up the *self-interests* in S1 (also highlighted in yellow). *Equity considerations* are repeated in S2, S3 and S4 but from different perspectives. Moreover, *unlikely … will*

Table 6.3: The role of phraseology in identifying the writers' position and citation function in Example 6.8

Text	Signals of topic/argument	Signals of citation/attribution	Signals of contrast/continuation
Negotiations about the initial distribution of permits will**, of course,** be dominated by the **self-interest** of participating nations	*Negotiations, self-interests, distribution*	*n/a*	*n/a*
***Nevertheless**, **it is often maintained** that considerations of international **equity** will also play some part, even if only as providing 'focal variables', to use Barrett's term, around which the negotiations can be centred..*[fn]	*Considerations of international equity*	*it is often maintained*	*Nevertheless*
***Indeed**, explicit **recognition** of the role of **'equity'** considerations **is contained in** the Framework Convention on Climate Change agreed at the 1992 Rio de Janeiro World Environmental Conference (INC 1992).*[fn]	*recognition of the role of equity*	*is contained in*	*Indeed (continuation)*
However, the widely **conflicting interests** of different countries as regards the extent and form of measures to restrict carbon emissions **makes it unlikely that genuine equity considerations** will play **any role** at all as compared with naked bargaining, as the March/April 1995 Berlin conference on the subject has demonstrated.	*Conflicting interests,*	*n/a*	*However*

play any role at all picks up *will also play some part* in S2 in the sense that the writers in S4 are denying any role of equity considerations. In addition to the use of the cohesive devices (e.g. *Nevertheless*, *Indeed* and *However*), the role of phraseology is important in reflecting the extent to which the writers make their arguments explicit in the block. The lexical repetition in this block reminds the reader of previous arguments in the same block, for easier understanding of citation functions.

Taken at the clause-level, S2 only *acknowledges* the source of the term '*focal variables*' (Barrett) and also presents Barrett as one who believes, among others, that equity considerations will play some part in the negotiations process. At the block-level, however, S2 is a multi-functional citation in that it Signposts the readers to the source of the mentioned term (Barrett) as a particular example of people who maintain this view on international equity. This function is achieved in the side note (referred to by ^{fn}), where the references are given, via the use of the imperative verb *See* (*See, in particular, S Barrett, …*) which specifies Barrett as one example of many people who maintain that *considerations of international equity will play some part …*. The second and most important function in the block is Engage (defined in this study as expressing critical dialogue with the source' view). While the writers in S1 insist that the negotiations about TP distribution will be affected by the self-interests of the participating countries, the source in S2 is cited in a way which shows opposition to (Engagement with) the averred statement in S1. Harwood (2009) notes that Engaging citations are often multi-functional: S2 is one example.

At the block-level, Example 6.8 is clearly marked for explicit argumentation. The use of disjuncts (e.g. *However*, *Indeed* and *Nevertheless*) to signal the relations between clauses, statements or paragraphs makes the argument explicit. This is consistent with the finding from Thompson et al. (2017) that papers in constellation 5 score highly on explicit argumentation. The majority of examples in this paper include explicit signals of the links between propositions. In general, the relations between the first four statements of Example 6.8 can be summarised in Figure 6.2:

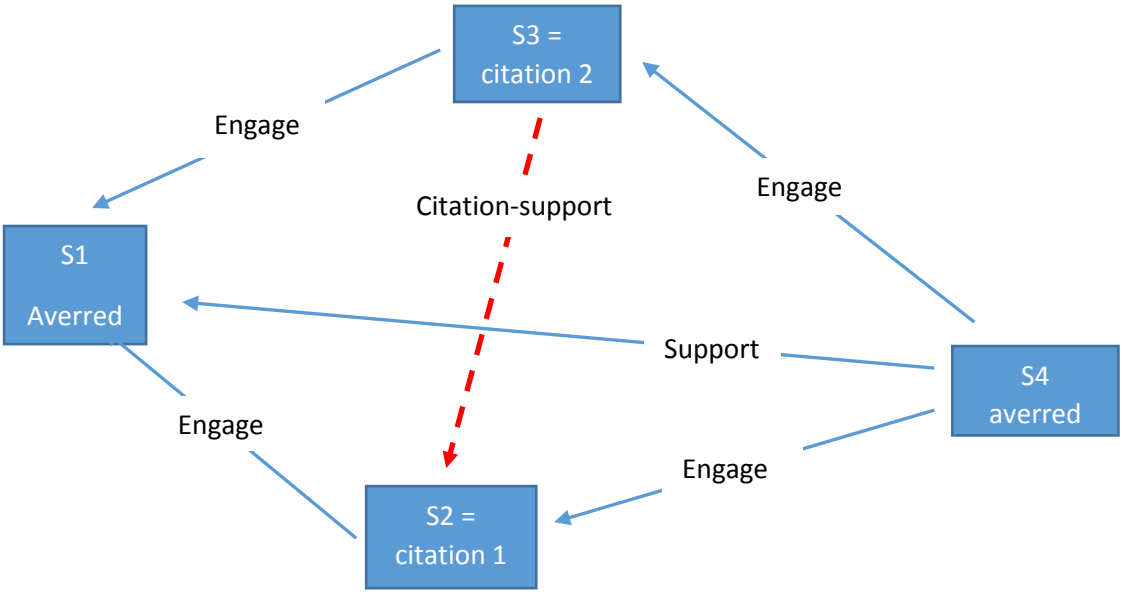

Figure 6.2: Visual representation of the cited-citing relation (citation function) and the cited-cited relations at the block-level in the first four statements of Example 6.8

Two more examples (6.9 and 6.10) in this regards include elements of the complex interaction in the citations of this paper. As the analysis of the two examples unfolds, it should be easier to identify the citation functions and the writers' position at the clause- and block-levels. In Example 6.9 below, the block contains two citations (in S1 and S10). In S1, the writers *acknowledge* the sources which have described one cited rule as *Kantian*. To describe something is a neutral act which involves no stance unless stance is marked by the presence of an accompanying evaluative lexis such as an adjective or adverb (e.g. has been *effectively* described). Some sources which have described the proposed rule as Kantian are listed in the respective side citation.

The writers introduce this description not only to Signpost readers but mainly to Engage with the sources' description of this rule as Kantian. Moreover, consider S8 and S10 and how they relate to S1. In the second paragraph (S8), the writers suggest that the description of the cited rule as *Kantian* is inaccurate (e.g. *However, there is nothing Kantian about this rule at all*). Therefore, S8 contradicts the description and explanation of this rule as *Kantian* in S1, S2, S5 and S6. S10 shows that the sources in S1 have misunderstood Kant's description and argument. Unlike the use of insist in Example 6.7 (*other commentators, too, have insisted* ...), insist in this example is presented in Kant's defence (i.e. that he has made his argument clear and that the sources in S1 have misunderstood the Kantian approach).

Example 6.9:

> [1] *One proposed rule* ***has been described as*** *a 'Kantian' allocation rule.*[fn] [2] **This** rule is that all countries are asked to say what uniform reduction in emissions they would prefer. [3] It is expected that countries will give different answers, of course, depending on how they perceive the benefits they will obtain from general uniform abatement. [4] **Hence**, there will be no agreed uniform reduction. [5] **But** it is proposed that each country should then reduce its emissions by the amount of uniform reduction it had originally proposed. [6] **For example**, China might propose that everybody reduce emissions by only 0.2%, **since** that corresponds to its own interests and the USA might propose that everybody reduce emissions by 4%. [7] If no agreement is reached China would have to stick to its proposed 0.2% reduction and the USA to its proposed 4% reduction.
>
> [8] **However, there is nothing Kantian about this rule at all**. [9] **For the essence of Kant's** 'categorical imperative' is that, in arriving at a principle to govern everybody's behaviour, we should all abstract ourselves completely from our own particular interests and propose a uniform rule that we believe gives equal weight to the interests of everybody else. [10] ***Kant could hardly be more emphatic in insisting that*** *what is morally right cannot be determined by what lies in one's interests.*[fn] [11] **This means, in the present context, that if** all countries had equal information about everybody else's interests, they would all propose the same uniform reduction in emissions.
>
> [12] A system under which each country proposes, instead, what uniform reduction happens to be in its own interest bears no relation at all to the Kantian requirement of impartiality. [13] It would be like a rule emerging from a process in the course of which poor people opted for a rule under which all rich people were to transfer money to poor people and poor people were entitled to steal from richer people, and the rich people would opt for a rule under which rich people had no moral obligation to transfer anything to poor people (though they might do so out of charity). [14] If, then, no agreement is reached the rich people would be entitled to transfer very little and the poor people would be entitled to steal from them. [15] This is not at all what Kant had in mind!
>
> [16] **It is true that a Kantian approach would imply** - as do most egalitarian theories - that there is a moral obligation to ensure that priority was given to assisting the worst-off sections of the world's population. [17] **But this, too, does not imply transfers to governments as distinct from the poorer members of the world community**.

The writers present Kant's view in S9, S10 and S11. They go on to Engage with the implications of Kant's argument in the last statement in this block (S17). It could be argued that while S16 is conceded, S17 is asserted. In S16, *It is true that* is a marker of concession. In S17, the insertion of *But* in Theme position indicates Engagement with the implications of Kant's view in S16.

In bold are some cohesive markers (e.g.*Hence, But, For example, since, However, For the essence, This means* and *too*). Most of these cohesive devices occur in sentence-initial position. Such devices are viewed as signals of explicit argumentation in Thompson et al. (2017) in the sense that they show the extent to which writers choose to guide their readers through the arguments. Similarly, with a focus on citation, I argue

that such markers facilitate the identification of the writers' overall position and the citation function(s) at the block-level and also show the relationship between clauses and propositions.

Another example showing the complexity of dialogue between the interlocutors is Example 6.10. This example contains one citation (in S7). In this block, the writers introduce their alternative solutions which they believe should replace the cited principles and/or rules. This block contains the last citation in Beckerman and Pasek's paper.

Example 6. 10:

> [1] At the same time, the principle of 'capacity to pay', which is generally accepted in taxation theory as an equitable manner of sharing out the burden of financing collective goods and services, **should also be applied**. [2] **Thus although**, *cet par*, countries with high levels of carbon emissions per head would receive more permits per head, countries with high incomes per head would receive less permits per head. [3] The simplest formula that would take account of both these criteria, is one that gives most permits, per head, to countries with the highest ratio of current emissions (measured in any physical units) to GNP (measured in money at some agreed exchange rates). [4] **Such a rule would, of course, discriminate in favour of poor countries** since a lower per capita GNP (cet par) means a higher ratio of emissions to GNP and hence a higher initial allocation of TPs per head. [5] **It is true that if** poor countries also happened to be disproportionately low emitters they will also get fewer permits per head, but this would be perfectly 'fair' since they would presumably incur the least loss of welfare in reducing emissions to the permitted level. [6] **Anyway, poorer countries do not, on the whole, have disproportionately smaller levels of carbon emissions**. [7] ***For example, if this rule*** *were applied* ***on the basis of*** *1987 levels of GNP and carbon emissions, China would be entitled to six times as many permits per head as the USA, and India to twice as many.*[fn]
>
> [8] **However, it might still be argued that the above rule is insufficiently egalitarian**.[29] [29 is not a citation, it refers to a side note] [9] **But this can be handled by varying the precise mathematical formula to be used**, which could, **of course**, be a matter of **negotiation together with numerous other issues, such as** how many permits would be handed over to the international agency and the rules for the use of the revenues; or the extent to which some countries have already incurred heavy costs in order to reduce their CO_2 emissions (eg France); or the overall limit on emissions; or the time period over which 'current emissions' should be measured; or how to take account of future changes in emissions and GNP; and so on.

In S1-S5, the writers introduce one of the solutions they propose for better allocation of TPs. S4 and S5 include elements of concession (e.g. *of course* and *It is true that*, respectively). S6 contradicts (*Anyway*) the conceded statement in S5. The writers present a hypothetical example in S5 and contradict it in S6. S6 shows that the arguments in S4 and S5 are in some way hypothetical and possibly contrived. S7 is provided as evidence (or example) of the argument in S6, which could be seen as a

counter-assertion against the argument in S5. In other words, S7 Supports the argument in S6 indicating that if the arguments in S4 and S5 are true, then the distribution of permits would have happened in a certain way, based on the data obtained from the cited source. At the clause-level of S7, the writers simply *acknowledge* the source of data (which is the last column of Table 2 in the cited source: Barrett 1992).

In terms of the citation function, this citation fulfils different functions in the citing block, particularly the two contradictory functions Support and Engage. This Citation Supports, via exemplification (*For example*), its preceding statement: S6, which in turn engages with S4 and S5. Therefore, it could be argued that S7 Supports the writers' argument in S6, to challenge the hypothetical assumptions in S4 and S5. On the other hand, the same citation shows Engagement with the following text. For example, S8 shows that the writers enter into a critical dialogue (Engage) with the information in the previous statements, including S7. This Engagement is achieved by *However* which signals a reversal in the direction of argumentation.

In order to minimise the effect of the Engagement, the writers hedge the counter-argument (*It might still be argued*). This act of "mild" Engagement (Harwood 2009: 506) is relatively less marked than other examples (e.g. in S2 of Example 6.12 below "*ought to be rejected*"). Engagement in this thesis is inclusive of total or partial rejection or criticism of cited views. In Example 6.10, Engagement means that this rule, including the hypothetical example in S7, is still open to criticism by other researchers or the potential reader (compare Example 4.41 in Chapter 4: *O'Collins points out that … However, most island societies view …*). Engagement in Examples 4.41 and 6.10 is achieved by the writers (*However*) as well as by other people (*most island societies* as in Example 4.40) or by the writers and/or other authors or readers (*it might still be argued that* as in S8 in Example 6.10). S8 is heteroglossic as the writers acknowledge potentially different views; thus Engaging with the preceding statements.

In short, the first five statements in this block introduce the topic (the capacity to pay rule). The writers provide a hypothetical example in S5. S6 cuts the assumptions and possibilities and brings the argument back to the real world. S7 is an exemplar of the averred argument in S6 and provides a hypothetical example to Support the argument in S6. The additive discourse marker *For example* signals correspondence between what is yet to come and what has been just mentioned. In implicit encapsulation (as in

Rego 2001), such discourse markers act retrospectively as "implicit encapsulators" (ibid: 97) and have a vital function "in the construction of the argumentation in the text" as they guide the readers retrospectively as well as prospectively (Mosegaard 1998, cited in Rego 2001: 98). Moreover, the demonstrative *this* in S7 highlights this retrospective relationship as this citation suggests that the interpretation of this statement is dependent on an entity mentioned previously in the text. Example 6.10 includes explicit indices of the links between propositions (e.g. *for example, Anyway, However, But*) as well as implicit indices (e.g. *this*).

So far I have discussed the blocks as though they are discrete. However, the analysis also shows that the blocks are not totally discrete from one another. On occasions, additional information about the writers' overall stance and/or citation function is obtained by looking at two blocks together. There are four blocks in this paper which are closely connected. Due to space restrictions, only two of the blocks are described in this section (as Examples 6.11 and 6.12).

Example 6.11:

> [1] *One rule that* ***has been alleged to have some ethical force*** *is a rule which penalizes heavy emitters of carbon by allocating permits in inverse proportion to their per capita consumption of fossil fuels.****fn*** [2] The ethical principle behind this rule **would appear to be** some principle of 'retributive justice'. [3] **But, if we exclude a revenge motive**, **the only conceivable** ethical **justification** for such a rule would be that it constituted a disincentive for people to behave in an antisocial manner. [4] **But whatever the allocation of permits, as long as they are tradable** they provide the incentive to restrict one's emissions to the socially optimal level, **irrespective of one's initial level of emissions**.

Example 6.12:

> [1] *There seems* ***even less case*** *in terms of any ethical theory for rules that take account of how far countries have emitted carbon in the past.****fn*** [2] Apart from the **practical difficulties of estimation**, **such 'historical responsibility' rules and their related concept of 'natural debt' ought to be rejected on ethical grounds**. [3] The notion of moral responsibility is closely linked to the notion of being a free agent, voluntarily carrying out any act, in knowledge of its consequences. [4] The element of choice is an essential ingredient of moral responsibility. [5] But it is difficult to see how the inhabitants of any country can have had any choice as regards the amount of fossil fuels burnt up by earlier inhabitants, even if they are not immigrants. [6] It is difficult therefore to see why the existing citizens of any country should have to bear a heavier burden on account of their predecessors' consumption of fossil fuels.

The first block is given as Example 6.11 above. The citation in this block appears under a new cited rule (namely: 'Redistributive justice'). In this example, the citation in S1 is

distanced. This *distancing* is manifested in the use of the verb phrase *has been alleged to have* in S1. The verb *alleged* is equivalent to the verb *claimed* (identified by Martin and White 2005 as an indicator of *distancing*). In S3, the writers Engage with the cited rule and its proponents. The first sign of Engagement is *But*, the writers also indicate that this rule has a *revenge motive*. The way in which the writers present this justification (as *the only conceivable ethical justification*) suggests that this rule suffers from other issues (e.g. *revenge motive*). In S4, the writers Engage with the ethical justification of the rule (which was mentioned in S3) and this Engagement is signalled by the adversative disjunct *But*. S4 provides an alternative solution to the *per capita consumption* issue. Overall, the writers are critical of the cited rule, which is mentioned in S1. This block exemplifies the *distance*-to-Engage unconventional combination.

Continuing the criticism of this rule 'Redistributive justice', the writers negatively introduce the next rule: 'Historical responsibility and 'natural debt'' in Example 6.12. In fact, the last two citations (in Examples 6.11 and 6.12) can be combined in one block. However, they are divided into two blocks since they appear under two different sub-headings.

Example 6.12 represents one of the few blocks with a conventional combination of citation stance and function in this paper. The writers in S1 *contest*-to-Engage, in critical dialogue, the proponent of this rule. The use of the comparative expression *even less case* links this citation, which represents a new rule, to the previous rule in Example 6.11 and so provides an anaphoric reference to the previous rule. Moreover, *even less* is used here following the negatively evaluated rule in order to emphasise that the negative view applies even more to the new rule. In other words, the way this rule is cited means that it is less convincing for the writers than the previous rule and that the writers wish to position the readers in a way to be sceptical of the new rule. S2 is a direct criticism of the cited rule in S1 and therefore shows that this rule is included in critical dialogue with the writers' views (i.e. that it is inconsistent with the writers' view). The cited rule is criticised for issues related to *difficulties of estimation*. The writers also reject this rule based on ethical grounds. The writers explain these ethical grounds in the remainder of the block (see S5 and S6 in particular).

Because of the way in which the writers of this paper have chosen to construe their research environment, in opposition to other sources, one might expect to find many

cases of *contest*-to-Engage citations. However, this paper has shown preference for criticism of the sources beyond the clause-level of the citing statements. Chapter 5 has shown a tendency amongst the authors of this paper to *acknowledge* at the clause-level and Engage at the block-level. *Contest*, as a category of stance is uncommon in this paper. It seems that the writers prefer to neutrally present the cited arguments at the clause-level, but to achieve a number of functions, the key function is to show the writers' Engagement with the sources (i.e. that the source's arguments are inconsistent with those of the writers).

By a way of conclusion, it could be argued that although this paper contains 28 citations only, it exhibits how complex and important the relationship between averral and attribution can be. When Engaging with external research, writers seem to draw more on resources from averral and attribution than when they use external research to Signpost readers to that research or when they indicate a source's Position. The reason is that writers need to, and often do, shift between text averral and attribution more often when they are involved in critical dialogue with external views. This can be best understood in the citation block. The extensive and recurrent criticism of the sources makes this paper distinct from the remaining papers.

Criticism of other research is common in academic discourse, particularly in policy discussion articles. What is unexpected however, is the preference to do so beyond the clause-level of the citing statements. Beckerman and Pasek's criticism of the research they cite is not apparent at the clause-level of the citing statements. This is different from the other papers in the other constellation, and even from papers in the same constellation and time period (e.g. Kempton 1991; see Figure 5.9). For example, there are more *contesting* citations in Kempton (31.76%) than in Beckerman and Pasek (6.45%), although the latter paper tends to cite research with which the writers disagree.

Some classification schemes which identify the citation functions based on sentence-level analysis of citations may attribute the low occurrence of negative citations to an intention for Beckerman and Pasek to hide "their negative feelings behind positive or neutral words" (MacRoberts and MacRoberts 1984; cited in White 2004: 101). However, the block-level analysis of citations in this paper has clearly shown the writers to be explicitly critical of the research they cite. This finding demonstrates the importance of the block in reflecting the writers' overall position towards the research

they cite and the role (function) of that research in the citing text. Beckerman and Pasek do not hide their criticism of the research they cite; rather, they defer it until the citing statement unfolds in full.

Confirming this observation is the dominance of *Acknowledge*, which is the most frequent stance type (54.83%). Figures 5.9 and 5.14 show the *acknowledge*-to-Engage combination to be the preferred pattern in this paper. In fact, out of the 28 citations, 15 citations exemplify the *acknowledge*-to-Engage combination. A corollary of this finding is the low occurrence of *contest*, which is the most straightforward method of expressing negativity towards cited materials. Instead of direct involvement in critique with the (cited) authors at the clause-level of the citing statements, the (citing) writers prefer to Engage in a more sophisticated way with the cited theories or rules for TP distribution. This explicit Engagement tend to be identified outside the reporting statement (at the level of the block).

Overall, it seems that the authors of this paper are interested in opposing research and exploit it in a way to give rise to their arguments. This is achieved by weakening the sources' views or research in the citation blocks. In Example 6.6, for example, whilst *acknowledging* Barrett's criteria at the clause-level of citation, the writers show Barrett's criteria as misrepresenting the previously mentioned theories and relying on *ad hoc principles*. According to the writers, this makes them unable to *stand up well to scrutiny*. All the examples in this paper demonstrate the importance of the citation block in understanding the role of citation in the citing text and more importantly how the block-level evaluation positions the reader (e.g. to accept or reject cited arguments).

As noted above, the ways in which the authors of this paper present the cited research show the complexity of dialogue between the interlocutors (i.e. the writers, the cited authors and the reader). As an example, I draw on the signals of attribution and stance in Example 6.8. In that example, the writers concede the averred argument in S1 (*of course*). This concession is in dialogue with the reader and so it is a marker of the writer-reader interaction. At the same time, the second statement which includes the first citation (*Nevertheless, it is often maintained*) seems to be in dialogue with another author. In terms of the writer-author evaluation (Status; see Hunston 2000, 2011)[3], *it is often maintained* is a neutral signal of attribution. However, when combined with *However* and considered in the context of S1 (and S4) it becomes clear that this

attribution is included to Engage in critical dialogue with the writers' argument in S1. Therefore, it could be argued that this attribution signals the third way of interaction in this example (the author-writer interaction). The cited argument is presented as being in critical dialogue with the writers' view in S1 and then is supported, through another attribution in S3. S4 plays the most important role in this block; that is, showing the writers' negative evaluation of (Engagement with) the sources' views and positioning the reader to be critical of the sources' views in S2 and S3.

For this reason, the dialogic power of the Engagement system (Martin and White 2005) is manifested in such examples from this paper. Simply put, the writer-reader interaction in the citations of this paper is arguably more important than the writer-author evaluation. The writers tend to *acknowledge* the sources' arguments at the clause-level of each citation to achieve different functions. The main function is to position the readers to be critical of the sources' arguments, with which the writers Engage in the citation blocks.

6.4 Conclusion

Chapters 5 and 6 provide quantitative and qualitative evidence of variation in the citation practices between the two constellations and time periods. In particular, the current chapter investigates the variation between two papers from the two constellations. The selection of the two papers is the result of the observed variation in the occurrence and spread of the form, stance and function categories between the two papers as observed in Chapter 5. The figures in Chapter 5 showcase one paper from constellation 5 (namely, Beckerman and Pasek 1995) to be markedly distinct from the other papers. The same figures also show this paper to be particularly different from another paper from constellation 1 (namely, McAlpine et al. 2009).

The current chapter uses a text-based, close-reading approach to investigate the citation practices of the two papers. In terms of methodology, the analysis demonstrates that the block-level analysis is more important than that of the clause in the identification of, and comparison between, the stance and function categories. The way in which the analyses of citation form, stance and function combine demonstrates (particularly in the unconventional combinations) that the clause-level analysis of citation stance and/or function is insufficient and can on occasions be misleading. The results obtained from the clause- and block-level analyses cast doubt on the findings of some citation

classification typologies which classify citation functions based on analysis of citations in reporting statements only (e.g. Petrić 2007) or those which derive citation functions based on informants' responses (e.g. Petrić and Harwood 2013). In essence, the clause-level analysis presumes that all citations are conventional since stance and function are identified at the same level (e.g. Petrić 2007; Harwood 2009; Petrić and Harwood 2013). However, the combination of the two levels has clearly shown that citations are used in more sophisticated ways. In fact, if all citations were conventional (i.e. showing correspondence between stance and function categories as in *acknowledge*-to-Signpost patterns), then there would be no role for the block in the analysis of citation functions. As shown in Chapters 4 and 6, many citations are unconventional and perform different roles at the clause- and block-levels. In order to emphasise and understand their role, one has to identify and compare the stance and function of each citation at the clause- and block-levels. This comparison facilitates a fine-grained understanding of their role in the citing text.

Another point emphasising the importance of the block is that most of the writers' critique, for example in Beckerman and Pasek (1995), is carried out at the block-level, thus justifying the higher proportion of the function Engage. Although the writers of this paper show a preference for integral citations which facilitate the expression of stance, most of their critical dialogue with the sources is not prominent at the clause-level, which justifies the low occurrence of *contest*, the most straightforward way of expressing negativity towards cited research.

Based on the findings from the IDRD project (Thompson et al. 2017) that argumentation in constellation 1 papers tends to be implicit while in constellation 5 it is mostly explicit, this chapter provides evidence of explicitness in argumentation within the selected papers from constellation 5 (e.g. in Beckerman and Pasek) and implicitness in constellation 1 papers (e.g. in McAlpine et al.), but from the perspective of citation analysis. In this thesis, explicitness in argumentation is analysed at the block-level to identify links between clauses, sentences and paragraphs and to identify the writers' position towards the research they cite. In terms of stance, explicitness here relates to the writers' concern, or lack thereof (cf. Thompson et al. 2017), to make their evaluation explicit at the clause-level.

The two papers analysed in this chapter could both be classified as policy discussion papers. Yet, they show great variation in terms of citation practices. This finding gives prominence to the constellation approach for classifying the GEC papers (as introduced by Thompson et al. 2017). In other words, the constellation approach has proven useful as a new approach to the classification of the GEC papers. For example, this approach classifies the second paper (Beckerman and Pasek) as a social scientific paper. The authors discuss the various policies and approaches to the distribution of TPs from a purely political economist perspective. The analysis of citation practices in this paper confirms the findings of Thompson et al. (2017) about the characteristics of the constellation 5 papers. Even though this paper is distinct from the other papers in terms of the extent to which its authors are critical of the research they cite, it still exhibits some features of writing pertinent to social scientific writing and constellation 5 papers, such as focus on systems (e.g. the distribution of TPs), person-centred methods (e.g. Rawls' theory, modified Rawlsianism) and explicit argumentation (e.g. *Negotiations about .... Nevertheless, it is often maintained .... Indeed, explicit recognition .... However, ... Furthermore, ... But, as is argued below, this assumption …*; see Example 6.8). The authors of this paper discuss social views related to environment-related issues and their citation practices reflect a contingency on argumentation (e.g. *Kant could hardly be more emphatic in insisting that …*) rather than on universally-shared facts.

On the other hand, McAlpine et al.'s paper, whilst being classified as a policy discussion paper, shows variation in the citation practices from the previous paper. The authors of this paper approach the research they cite from the perspective of ecology, as a branch of life science. As their citations show, they seem to prioritise "the physical world over the world of ideas" (Thompson et al. 2017: 19). Their citations also reflect a focus on space and site of interaction between humans and the environment (e.g. Queensland, Columbia and Brazil), focus on measuring the quantitative impact of environmental concerns (e.g. *World meat consumption increased from 47 million tonnes in 1950 to 260 million tonnes in 2005 … (REF)*) and a focus on "abstracted human action" (ibid: 20) as in *The demand of these new consumers …, The impact of humans …* and *World meat consumption….* Their argumentation, in relation to citations, is contingent on facts (e.g. the increasing consumption of beef) and tend to be implicit (i.e. the writers tend not to make explicit links between propositions).

In summation, the two papers show variation in citation practices even though they both discuss existing policies and appear in the same journal. It could be argued that the textual comparison between the two papers reflect the interdisciplinary nature of the journal GEC, as receiving contributions from various disciplines (e.g. from political economy as a Soft Pure domain and from environmental ecology as an interdisciplinary branch of life and earth sciences; Becher and Trowler 2001). In general, the citation literature tends to compare citation practices between two or more disciplines, usually reporting similarity or variation between the selected disciplines. The current thesis, however, demonstrates variation amongst papers within the same journal, which implies that the journal is so far successful in maintaining its interdisciplinary appeal (as reported in Thompson et al. 2017). The identified variation may also be attributed to the different approach taken by each paper to the environment-related issue under discussion. The first paper discusses social policies, assorted case studies and highlights different locations where there is an increased consumption of beef. The same paper also discusses various solutions to mitigate the effects of increased consumption. The second paper Engages with the existing theories and rules for TP distribution.

The textual analysis of Beckerman and Pasek's paper (1995) and the comparison with McAlpine et al. (2009) paper confirm some of the quantitative findings from Chapter 5. For example, in the previous chapter it was found that most of the citations in the constellation 5 papers carry one stance type (see Figure 5.16). On the contrary, the same figures show that in constellation 1, most of the citations carry explicit and implicit stances. The examples in Beckerman and Pasek's and McAlpine et al.'s papers provide textual evidence of the internal variation between the constellation 1 and 5 papers.

In general, the implication of this chapter is that Beckerman and Pasek's (1995) paper differs from the other papers because of the way Beckerman and Pasek present and construe their research environment. The authors of this paper tend to cite studies to which they are antagonistic. The evidence of this argument is that Engage is the most frequent function in this paper only. On the other hand, the other papers with Support as the most frequent function (e.g. McAlpine et al. 2009) tend to cite studies with which the writers are in accord, and which at the same time Support their research. The findings from Chapters 4, 5 and 6 and the findings from the IDRD project are discussed in the next chapter with specific attention given to the implications of this research for

pedagogy, citation theory and future research in corpus linguistics and library and information research.

Notes

1- Compare this example to Example 4.17 (*One of the most respected climate modellers … argued*). In 4.17, the writers *endorse* the source while *acknowledging* their argument (*argued*). In that example, it was argued that *endorsement* of the source applies to their argument because it is in line with the writers' *endorsement* of the *prominent climate scientists*, which includes the source in S2, at the block-level (see S1 in Example 4.17). In Example 6.7, however, the case is different. While the writers describe the source as an *expert* and *acknowledge* his argument, the writers Engage with the source in the next paragraph (starting with S6). This Engagement shows that the writers do not agree with the source's argument in S4.

2- Example 6.8 appears in the previous chapter as Example 5.1. However, in this chapter, the complexity of the writer-reader interaction is highlighted in such examples.

3- For an extensive review, see Hunston's (1993, 1995, 2000, 2011) discussion of the relationship between attribution and evaluation. Hunston argues that evaluation is a key function of citation. In various publications (e.g. 1993, 1995, 2011) she explores the correlation between evaluation and attribution and whilst she acknowledges that a major function of citation is to evaluate sources, she investigates various ways of modifying attributions. One of which is the use of verbs or nouns of attribution.

Chapter Seven

Discussion of the Findings

7.1 Introduction

As stated in Chapters 1 and 3, this study is primarily driven by the need to better understand 'how citations act and are acted upon, in the citing text'. Another motivation for this piece of research is the perceived scepticism regarding citation indices and their ability to reflect citation value. The proposed model (Chapter 3) has been reformulated in a manner that facilitates a multi-level analysis of citations to account for their commonalities as well as the variations related to the way in which citations operate in different ways concurrently. Consequently, a large portion of this chapter is devoted to the discussion of the implications and applications of the proposed model. At the primary level, this chapter seeks to answer three questions:

1- What is the relationship between the citation form, writer's stance towards cited research and the citation function (its role in the citing text) and what the implications are as a whole for the study of citation?
2- How similar/different are constellations 1 and 5 between the time periods 1990-95 and 2008-10? And how do these findings relate to the character of GEC as an interdisciplinary journal? Are these findings anticipated, given what is currently known about GEC?
3- What are the implications of this research for a) pedagogy, b) corpus linguistics and c) Library and Information Science (LIS)?

The next section (7.2) answers the first question about the relationship between the three facets of the proposed model (citation form, stance and function) and the implications for future investigations of citation practices within and/or beyond disciplines, genres and other fields of knowledge. Then, Section 7.3 revisits the findings from Chapter 5 about the variation between the four data sets. This section also discusses the findings in relation to the IDRD project (Thompson et al. 2017), providing qualitative evidence of variation in citation between the selected papers from constellations 1 and 5. The following section (7.4) is further divided into two sub-sections: the first sub-section (7.4.1) briefly introduces the pedagogical implications of

the proposed model and the second sub-section (7.4.2) discusses the implications for research in corpus linguistics and LIS.

7.2 Discussion of the findings (Question 1)

The analysis of citation form, stance and function has resulted in the identification of conventional and unconventional combinations. In the former, there is a correlation between the identified stance and function categories (e.g. in the *endorse*-to-Support combination). In the latter however, the identified stance and function categories do not match (e.g. in the *distance*-to-Engage combination). In the unconventional combinations, it is possible for any function to co-exist with a different stance type (e.g. Engage as a function can be identified at the block-level to co-exist with *endorse, acknowledge* or *distance* as stance types at the clause-level). The combination of stance and function analyses in the present model adds to the theory of citation as it facilitates textual analysis of how citations are presented at different levels and their relevance to the citing text.

The combination of the different aspects demonstrates the importance of the model in reflecting the complexity of citation in interdisciplinary discourse. The constellations approach (as proposed by Thompson et al. 2017) mimics, to some extent, the different disciplinary origins of the papers in GEC. I would argue that the proposed model has been successful in revealing variation between the two most distinct constellations. One of the main aims of this thesis was to build a model that is exhaustive, yet concise, in order to account for the basic roles performed by citations and to make the model applicable to other research contexts. Future investigation of citation practices should perceive citation as a bi-directional phenomenon that is affected by, and also affecting, the citing text as well as the reader.

In the citation literature, there is a lack of research which combines the analysis of citation form, stance and function in order to investigate the role of citation in the citing text. Even when stance and function analyses are combined (e.g. Coffin 2009), both analyses are still conducted at the clause-level of the citing statements. In proposing a new level for analysing citation functions, the current thesis argues, and also demonstrates that citations operate at different levels and that the analysis of one level is insufficient. The block-level analysis of citation functions demonstrates, unlike most citation classification schemes, a tendency amongst some writers to evaluate cited

propositions beyond the reporting sentence. Any analysis of citation functions should take into account the importance of the surrounding context in revealing the different roles performed by citations, especially when those roles cannot be identified at the reporting sentence level. In summation, while this thesis distinguishes the three aspects of analysis, it shows that they are interconnected and that the choices within each affect knowledge construction and comprehension within and beyond the level of the citing statement. For example, the integral citations are associated with more explicit stance whereas the non-integral forms allow writers to conceal their stance towards the cited propositions, which tends to be expressed more clearly beyond the citing sentence-level.

7.3 Discussion of findings continued (Question 2)

Overall, in the early years, there is divergence between the papers from constellations 1 and 5. This divergence occurs in the use and spread of citation form, stance and function categories. The early period of constellation 5 shows divergence from both periods of constellation 1 and the recent period of constellation 5. In the recent time period, it seems that the journal has established its own character and thus the two constellations have converged. It could be argued that the journal, in some way, bridged the gap between the two constellations, and its interdisciplinary nature could well be the reason behind the later convergence. There is no reason to suppose that the journal intentionally changed the way writers wrote, but in terms of citation practices, the convergence seen in the recent papers may be expected as a consequence of the journal becoming more focused.

What was unexpected however is the diachronic change in the constellation 5 papers, which represent the social scientific component of the corpus. This finding was unexpected, given that the GEC website explicitly states that the journal is interested in articles with "a significant social science component" (retrieved October 2017). My own expectation was that the journal as a whole, would change its focus from science-like papers (as in constellation 1) to social science-like papers (as in constellation 5). However, this has not been the case. Both constellations seem to have recently moved into an area in between, with the most observed change occurring in constellation 5. This could possibly be attributed to constellation 5 having changed its citation practices and engagement with sources. For example, the writers in the later period of

constellation 5 tend to construe their research environments differently from those in the early years of the same constellation. As noted in Chapter 5, there is less Engagement with sources in the later years of this constellation and more use of citation in Support of citing texts. This could well be the consequence of the journal establishing its own voice and citing style.

The value of the proposed model is demonstrated by its application to a research question about the nature of (inter-)disciplinary discourse. In other words, the proposed model has revealed variation in terms of citation practices in the selected GEC papers. As an interdisciplinary journal, GEC receives contributions from different disciplines. The proposed model has proven useful in demonstrating variation in citation practices between some papers which represent the scientific and socio-scientific components of the GEC corpus. This variation was expected based on the findings of the IDRD project (Thompson et al. 2017) about the nature of GEC as an interdisciplinary journal and, more importantly, about the two constellations reflecting greatest variation in their dimensional profiles. I would argue that the proposed model is successful in reflecting how citations act, and are acted upon, in citing texts based on evidence from a successful interdisciplinary journal such as GEC. As a result, the proposed model can be applied to other disciplines or genres in order to trace synchronic and diachronic variation within or between genres and/or disciplines.

7.4 Discussion of the findings continued (Question 3): implications for pedagogy, Corpus Linguistics (CL) and Library and Information Science (LIS)

This thesis has shown (see Chapters 4, 5 and 6) that published writers combine stance and function both conventionally and unconventionally. The identified combinations reveal useful implications for 1) pedagogy (7.4.1) and 2) CL and LIS (7.4.2).

7.4.1 Pedagogical implications

As for pedagogy, a substantial body of the literature has discussed the pedagogical implications resultant from citation analysis. In the last three decades or so, research has shown that non-native students and novice researchers display a lack of awareness of the ways to cite and the reasons for citing (e.g. Thompson and Ye 1991; Thompson

and Tribble 2001; Pecorari 2003, 2006; Abbasi et al. 2006; Chandrasoma et al. 2004; Errey 2007; Harwood 2010).

In the early 1990s, a large portion of citation research has stressed form over purpose of citing (e.g. Weissberg and Buker1990; Jordan 1992, cited by Thompson 2001: 197-8; Swales 1990). The implications of such research were discussed in terms of what each form tells the reader. For example, Weissberg and Buker (1990) identified integral citations as author-focused citations and non-integral citations as information-focused. Swales (1990) went on to investigate stance in reporting and how it can be realised in different forms of citing (see Chapter 2, Sub-section 2.3.2.2 for more detail).

Over the last twenty years or so, there has been an increasing awareness of the functions of citations and how they relate to citation forms (e.g. Thompson and Tribble 2001; Thompson 2001; Petrić 2007; Coffin 2009; Fardi et al. 2017).[1] More attention was given in this research to the purpose of citing at postgraduate (PG) level. The general conclusion from this research is that PG students and also novice researchers, struggle to understand how they should present external research and for what specific purposes. This lack of understanding results in the production of "anomalous and difficult texts by choosing the wrong kind of reporting sentence" (Shaw 1992: 303).

The difficulties and issues associated with reporting and citation have long been discussed in the literature as have been the proposed solutions to address them. In general, the previous literature has focused on issues associated with citation and reporting such as problems with quotations, paraphrasing and tense and aspect usage. More recently, there has also been an added focus on the lexical and grammatical choices that are available at the students' disposal when citing (see Chapter 2 for details). The previous literature discusses the implications and proposed solutions as predominantly those that apply to the teaching of reading, writing and grammar.

The general conclusion is that EFL teachers and EAP courses at some stage should "sensitise students to the [lexical and grammatical] choices that are available when citing (Thompson and Tribble 2001: 91; emphasis mine) via extensive explicit exposure to the practices of published writers or high-scoring PG students in order to allow less successful students to compare their citation practices with those of more competent writers (e.g. Petrić 2007; Dontcheva-Navratilova 2016). The issue with such research, however, is that it is confined to the clause-level of the citing statement. It is true that

students are still struggling to understand the nature and purpose of citing, as Harwood (2010) notes. This struggle is evident in their lexico-grammatical choices at the clause-level with respect to tense, reporting verbs and structures. However, these issues at the clause-level create other issues for the comprehension of citations at the block-level. In other words, improper presentation of cited propositions at the clause-level creates difficulties for the understanding of their role as well as the writer's position in a longer context. The previous literature tackles the issues associated with citing and presenting sources at the sentence-level, leaving out an important aspect in the production and comprehension of citation: that is the block.

While acknowledging the importance of tackling the issues identified at the clause-level of citation, the current thesis adopts a different approach to raising the students' metacognitive awareness of citation and reporting. Despite persistent attempts at improving student citation practices at the clause-level, research shows that students still display a lack of understanding of how to present sources at this level (e.g. Harwood 2010; Ali 2016). In essence, the students' lack of awareness and misrepresentation creates issues not only at the sentence-level, but also in the comprehension of their positions and citation roles in the texts they produce. Therefore, I would argue that it is useful to commence the training process in relation to citation with the block as the unit for the deconstruction and composition of citations. In fact, approaching citations at block-level enables a more exhaustive understanding of the choices within each clause and their consequence on the build-up and coherence of the writer's argument. Before discussing the implications of this approach, it is useful to note that the block has various implications for the teaching of reading, writing and grammar. I shall go on to propose that future research is undertaken with a view to designing courses and activities that tackle these issues in the block not only to make PG students aware of the significance of their choices at the clause- and block-levels, but also to show how citations act, and are acted upon, in the citing text.

In terms of reading, the current thesis argues that it is useful to start with raising the students' awareness of their own citation practices and how they compare to those of published writers. Students could be asked to reflect on their lexical and grammatical choices at the clause-level and beyond. Students should also be exposed to successful citation practices in order to show how their practices differ from others. This exercise has the benefit of not only showing the manner in which successful writers cite external

sources, but also the benefit of drawing the students' attention to what teachers and/or examiners look for in citations so that the students know which practices they should or should not engage in. The teachers could possibly help their students by setting reading exercises to identify what is lacking or being misrepresented in the students' written texts. The findings could also be compared with those who are more capable.

Using blocks in the teaching of reading has various implications and potential benefits for raising the students' awareness of 1) writers and authors positions, 2) how writers position their intended audience (i.e. either to accept, reject or be undecided about sources' views), 3) clause-level choices and the consequence on the block comprehension (e.g. tense, reporting verbs, cohesive devices and other stance markers), 4) block-level choices and their effect on the understanding of reporting sentences and their functions (i.e. signals of citation function such as those signalling change of position such as *However*), 5) the writer's concern or lack of it for making argumentation explicit or implicit (Thompson et al. 2017), 6) ways of expressing implicit stance (see Section 4.2 in Chapter 4) and 7) how the various categories of form, stance and function combine to produce conventional and unconventional patterns.

To illustrate how this may be achieved at the block-level, I draw on two examples. The first example appears in Chapter 6 (Example 6.8) where the writers invest more textual resources to make their argument explicit in the citations' block (this is an example of the constellation 5 papers). The second example is taken from a previous study but is included to show how writers position themselves, the cited authors and the putative audience. The writers in this example explicitly evaluate the authors at the clause-level of every citation. What is remarkable is that the writers in this example seem to be less concerned with explaining the steps involved in argumentation than the writers in the first example. These examples are presented in Tables 7.1 and 7.2 together with signals of topic, attribution or citation, writers' stance, authors' stance when applicable and signals of citations functions.

At the level of the clause, both examples have signals of topic, attribution, writers' stance (e.g. *maintained* in 7.1 and *skeptics, reality, significance, evidence, gone so far as to declare, paradoxically* and *claims* in 7.2) and authors' positions (e.g. recognition in 7.1 and *accused, overreacting, exaggeration, human desire for attention, biases them, deception* and *hoax* in 7.2).

Table 7.1: Example 6.8 revisited (represented as Example 7.1)

Text	Signals of topic/argument	Signals of citation/attribution	Signals of contrast/continuation of argument/position
Negotiations about the initial distribution of permits will**, of course,** be dominated by the **self-interest** of participating nations	Negotiations, self-interests and distribution	n/a	n/a
***Nevertheless**, **it is often maintained** that considerations of international **equity** will also play some part, even if only as providing 'focal variables', to use Barrett's term, around which the negotiations can be centred..*[fn]	considerations of equity will also play some part	it is often maintained	Nevertheless (contrast)
***Indeed**, explicit **recognition** of the role of **'equity'** considerations **is contained in** the Framework Convention on Climate Change agreed at the 1992 Rio de Janeiro World Environmental Conference (INC 1992).*[fn]	recognition of the role of equity	is contained in	Indeed (continuation)
However, the widely **conflicting interests** of different countries as regards the extent and form of measures to restrict carbon emissions **makes it unlikely that genuine equity considerations** will play **any role** at all as compared with naked bargaining, as the March/April 1995 Berlin conference on the subject has demonstrated.	conflicting interests, unlikely that equity considerations will play any role	n/a	However (contrast, revering to writers' position)

Table 7.2: Example 7.2 (taken from from Brysse et al. 2013)

	Citation	Signals of topic	Signals of citation/ attribution	Signals of writer's stance and position	Signals of authors' stance and position	Signals of citation function
S1	*Over the past two decades,* ***skeptics*** *of the* ***reality*** *and* ***significance*** *of anthropogenic climate change have frequently* ***accused*** *climate scientists of ''alarmism'': of over-interpreting or overreacting to* ***evidence*** *of human impacts on the climate system (e.g., REFs).*	reality and significance of anthropogenic climate change, alarmism	Skeptics … have frequently accused,	skeptics of the reality and significance, evidence of human impact	accused, alarmism, over-interpreting, overreacting,	
S2	***Often it is alleged*** *that the motivation for such* ***exaggeration*** *is to gain media attention and funding for research,* ***suggesting*** *that scientists' human* ***desire for attention*** *and practical need for funding* ***biases them*** *toward exaggerating threats (REFs).*		Often it is alleged that,	alleged, suggesting	Scientists' human desire for attention, biases them, exaggeration	
S3	***Some extreme skeptics have gone so far as to declare*** *global warming a ''****deception''*** *and even a ''****hoax''*** *(REFs).*		Some extreme skeptics have gone so far as to declare	extreme skeptics, gone so far as to declare	deception, haux	
S4	***Paradoxically****, since the release of the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC)* ***such claims*** *have become more frequent, even as the quantity, quality, and diversity of relevant* ***scientific information supporting anthropogenic climate change has vastly increased*** *(REFs).*		Such claims	Paradoxically, such claims		Paradoxically, even as

At the block-level, however, the writers in the first example make the steps involved in argumentation more explicit than the writers of the second example. The writers in 7.1 make the connection between clauses and paragraphs explicit, signalling change and/or continuation of the writers' and sources' arguments. Examples of these devices include *Nevertheless*, *Indeed* and *However*. These devices function to organise the discourse and more importantly indicate the writers' stance and citation functions at the block-level. In previous research (Aljabr 2011), I found that undergraduate university students in one Saudi university overuse additive cohesive devises when producing texts, such as *also* and *in addition*. Providing authentic examples such as 7.1 gives students the opportunity to see how such devices and others (e.g. adversative connectors such as

Nevertheless and *However*) can be managed in academic writing and so enables the students to balance their use and distribution of these cohesive links. It also raises their awareness of the importance of such devices as textual organisers and more importantly, indicators of the writers' position and citation functions.

The writers in the second example invest more at the clause-level of every citation to make their position as well as the sources' more explicit. With the exception of the anaphoric references in *such exaggeration* and *such claims*, the writers draw no links between the different attributions and their averrals in the block. In other words, this example could be used to teach students the interplay between writer and author positions and also the effect of the persistent negative evaluation of the authors' views. The writers are very critical of the sources' views in every citation, even though one argument prevails in this block: *reality and significance of anthropogenic climate change*. The writers in Examples 7.1 and 7.2 tend to cite research to which they are antagonistic. However, the writers in the second example make their stance towards the cited propositions/views explicit at the clause-level of every citation whereas the writers in the first example tend to express their negative stance at the block-level. Introducing activities which facilitate comparison between the citation practices of different texts, as in the previous two examples should enable students, with the support of teachers, to understand how the writers' choices at different levels influence the ways in which texts are comprehended and the extent to which writers make their argumentation explicit or implicit at the clause- and block-levels.

This practice also has useful implications for the teaching of grammar and tense choice indication. Pecorari (2013) points out that citation patterns are intricate and specific. She links her implications [as well as those of Shaw's (1992)] to Hoey's (2005) concept of lexical priming. Pecorari (ibid) argues that teaching information structure and linguistic realisations should enable students to eventually know which choices are perceived as appropriate and conventional (e.g. appropriate reporting structures, verbs and signals of citation/attribution). Extensive exposure to successful citation practices should enable students to acquire complex chains of citing and to realise the best lexicogrammatical choices for signalling citation, the writer's and author's stances and more importantly how these choices will affect the ways in which citations function in the citing texts. However, unlike Pecorari and other researchers, I would argue that this prolonged exposure has not been so successful in raising students' awareness towards

their lexico-grammatical choices at the clause-level and their effect on the understanding of their position or citation functions at the block-level.

EAP practitioners can also expose students to the various findings of tense-in-reporting studies (e.g. identified patterns of usage; see Table 2.1 in Sub-section 2.3.2.1). Taking the block as the unit for analysis, teachers and students can highlight the tenses used in reporting previous research. For example, the use of the past tense to report previous research is a marked choice which decreases proximity between the cited and citing texts (Swales 1990). Students are generally taught to use the simple past whenever they cite previous research with a specific date of publication. In academic writing, particularly in MA or PhD thesis writing, such choice is uncommon and signals to the reader a distance between the cited and citing texts (e.g. to show cited research to be old and no longer valid). At the clause-level, the use of both tenses is grammatically correct. However, when academic writers cite old and recent research, they might choose between the past tense to report old research and the present to report recent research. In this view, it could be argued that there are two competing views of tense usage in reporting. The traditional view links tense to time, at least in some EFL contexts. The other view (including Swales 1990) links tense to relevance or proximity with the citing texts. Only block-level analysis can illustrate the different uses of tense and also the consequence of the shift between tenses on the build-up of the writer's argument and posture change (Sinclair 1985). Compare Examples 2.11 and 2.12 in Chapter 2.

Students can also be exposed to the conventional and unconventional combinations of citation form, stance and function. Such activity would raise their awareness of the ways in which citations operate at different levels and in different directions in the citing texts (see Sub-sections 4.3.1 and 4.3.2 for examples). It will also make the distinction clear between stance and function analyses and how writers evaluate sources' contributions to achieve a variety of functions. This will sensitise students to the complexities associated with how writers position themselves with respect to the cited research, how they position the authors for the attention of the reader. Once the students achieve a certain degree of competence at this level, they should realise the ways in which citations are evaluated and used in citing texts, and eventually recognise different moves or steps that are involved or common in citation blocks.

Following extensive exposure to reading and grammar activities, teachers can then design writing tasks aimed at improving citation practices and concurrently assessing their students' achievement. Such tasks should be aimed at tackling the issues identified at the clause- and block-levels (e.g. tense, stance and function choices and how they affect the way in which citations relate to the citing text). Teachers can then interview students about their choices and modify the ways in which blocks (and also clauses) are written, to see if and how this affects the comprehension of citations and their role in the text. Once the students conform to the writing conventions of their institutions or unique fields of study at the clause- and block-levels, they should be able to produce more coherent texts, successfully locate their new research within existing literature and most importantly draw various relations between cited sources and their ongoing discussions.

It is important to note however, that mastering the above skills in terms of citation is a non-trivial task. This is because it is a long and time-consuming process. It also depends to a large extent on the individual differences between language learners and their ability to grasp the different ways of presenting, evaluating and employing citations. Yet, the current thesis has proposed a new unit for the deconstruction of citing texts. It is argued that approaching citation at this new level facilitates the identification of textual cues within and beyond the clause and their effect on the ways in which the text unfolds in its context of operation: the block.[2]

7.4.2 Implications for research in Corpus Linguistics (CL) and Library and Information Science (LIS)

This sub-section discusses the implications of the proposed model, together with the findings, for CL and LIS. This thesis has so far demonstrated the importance of identifying the citation block and how the clause- and block-level choices contribute to what should and should not be considered in function analysis of citations. In terms of CL, this sub-section seeks to answer the question of whether the proposed model and the findings may be generalised, or replicable in other research contexts (e.g. in other genres or disciplines).

In applying the proposed model, the current thesis has shown that the proposed model and its qualitative application could be generalised and applied beyond GEC papers. Citations perform different roles when analysed at clause- and block-levels. The

proposed model investigates writers' choices at every level to observe how citations are evaluated and employed within and beyond the citing statement. In terms of the quantitative findings, the current thesis provides evidence of synchronic and diachronic variation between some GEC papers from constellations 1 and 5, even though the reliability of the evidence is reduced by the small number of papers selected. The quantitative findings reflect tendencies in the selected papers only and may not be applicable to all GEC papers or interdisciplinary discourse. Due to the fine-grained nature of the analysis, it was decided to limit the number of selected papers used in the investigation of citation practices at the proposed levels and directions (see Chapter 3). The quantitative application of the model in Chapter 5 is included to indicate synchronic and diachronic variation between the four sets of data and to link the findings to those observed in the IDRD project (Thompson et al. 2017) with respect to the character of GEC and the difference between the two constellations. In summation, the current thesis provides qualitative and quantitative evidence of variation between citation practices in the selected papers drawn from the two constellations. The main aim of this thesis, as noted in Chapter 3, is to propose, not to quantify, a model that can better account for the role of citation in interdisciplinary discourse. The quantitative findings therefore, are included to give an indication of the variation in the use of the different form, stance and function categories.

The proposed model departs from the contention that citations are unequal and play different roles in citing texts. To account for their roles, the block was introduced as a new analytical unit for analysing citation functions. It is argued that citations operate in blocks and to understand their functions, it is necessary to investigate how the writer evaluates cited propositions within and beyond the clause-level of every citation. The investigation of the various form, stance and function categories in the 1186 citations resulted in the identification of conventional and unconventional patterns of citing. The ultimate focus within each pattern was on the functions performed by the respective citations. The accompanying stance categories could be taken as another resource for prioritising the different uses of citations in the unconventional combinations.

Before I discuss the importance of the combinations and the ways in which they can be linked to citation value, I shall argue that future investigations of citation practices should start with the block, rather than the clause, as the main unit for analysis. This has the benefit of identifying textual cues that contribute to the understanding of the

writers' position within and beyond the citation-level. In terms of corpus-based investigations of citation functions, I propose a new context for investigating citations, arguing that citations should be first divided into blocks. Every block may contain one or more citations. Analysts should therefore investigate key-citations-in-blocks (KCIBs). The block boundaries are likely to be different for every citation but in general they should not exceed two paragraphs before and two paragraphs after the citing statement for a number of reasons. The reasons have been enumerated in Chapter 3.

Once citations are divided into blocks, it would be easier for corpus linguists and/or discourse analysts to identify conventional and unconventional patterns of citing as identified in the proposed model. For example, using a medium-size corpus, the analyst should be able to analyse and code all the citations in their pre-identified potential blocks. The analyst may wish to pre-process the corpus and code every citation in terms of form, stance and function categories. While this process is time-consuming, it has the benefit of facilitating comparison between the clause- and block-level choices and their effect on the construction and understanding of citation. The identification of the different stance and function combinations gives a sense of what should and should not be counted in the analysis of citation role and its relevance to the citing texts. The block-level choices can be more important in reflecting the overall functions of citations than the choices at the clause-level. To illustrate this, I draw on one example from Flϕttum et al. (2006: 242):

Example 7.3:

> *Smith (2000)* ***claims*** *that phenomenon X is* ***difficult to describe*** *in linguistic terms*; **however**, in my studies (Johnson 2001) it is **essential** to the explanation of text structure.

Studying the presentation of other voices, Flϕttum et al. (2006) investigate the above example in terms of the role of negation and concession in discourse. They specifically argue that negation in such examples indicates "what the function of the voices integrated … is" (ibid: 243). I draw on the citation in the first clause of the above example (*Smith ... claims*) to emphasise how the writer's choices behind the clause-level in that example reflect the function of that citation in the block (which contains two clauses in the same citing statement). At the clause-level, the writer *distances* Smith's proposition. At the proposition-level, however, the same writer Engages with the first citation, arguing against Smith's claim and criticism of the *phenomenon X*. The

writer uses the adversative cohesive device *however* to indicate a change in argumentation and to introduce their position of the *phenomenon X*, as being *essential to the explanation of text structure* in their studies. In the first clause, the writer withdraws from responsibility for the attributed proposition. In the second clause, the same writer takes a more critical approach of Smith's view, thus contracting the space for reader support of Smith's claim. The use of *essential* reflects the overall position of the writer towards the *phenomenon X* and the function of the first citation. This citation exemplifies the *distance*-to-Engage unconventional combination.

Returning to what should and should not be considered in the analysis of citation functions, I argue that the block-level cues (*however* and *essential*) should be given prominence over the clause-level cues (*claims* and *difficult to describe*). That by no means implies that the clause-level resources play no role in function analysis; instead, it means that the block-level resources play a more vital role in the identification of the writer's overall position and thus the citation function(s).

The citation function(s) can then be identified in relation to the writer's argument/finding/discussion in the block. In terms of relevance, citation functions can be divided into internal and external in relation to the citing text. In the text-internal functions such as Support and Engage, citations play an important role in the construction and build-up of the writer's argument in the block. Regardless of whether this role is positive (in the function Support) or negative (in the function Engage), these functions can still be clearly linked to the citing text and therefore are by default more valuable than citations performing text-external functions such as Signpost and Position. The text-external functions simply direct readers to external research or other authors' positions (see Table 4.1 for examples) and as a result are not as valuable as the text-internal functions. They remain important in locating research within the existing literature, but are not directly relevant to the writer's argument in the block, as are the text-internal functions. They generally function to guide readers outside the citing text and therefore are of low value, compared to those which clearly relate to the block's argument. In other words, the cited author's contribution in Signposting or Positioning citations is less important than their contribution when it is directly relevant to the citing writer's argument.

Taking all that into consideration, it seems reasonable to argue that the model can be applied to a larger corpus, given that time and resources are allowed. However, as noted in Chapter 3, applying the model to a large corpus means there is less room for detailed qualitative investigations of citation practices. For this reason, it could be argued that future research in CL can develop or possibly use some corpus tools which are specifically designed to divide citation blocks and capture cue words such as reporting verbs, nouns, structures, cohesive devices and other stance markers. An alternative approach would be to allocate in a research group, different roles for the researchers involved in citation analysis. For example, one researcher divides or confirms the results of automatic identification of citation blocks. Another researcher codes the identified citations in terms of form, stance and function. Another researcher combines the identified stance and function categories for every citation in order to produce another list of conventional and unconventional combinations.

If resources allowed, it would be extremely useful to identify and classify citations in terms of form, stance and function. The classification would have to rely on linguistic cues within and beyond the clause-level, given that the writer's argument is explicit at both levels. However, one issue with automatic coding is the inability to distinguish between clause- and block-level choices and the extent to which they play a role in the identification of citation function. For example, in the previous example, corpus tools or automatic classifiers would combine all the stance and function markers in one group (e.g. *claims, difficult to describe, however* and *essential*). Without close-reading, it seems difficult for such corpus tools to identify which cue(s) is(are) more important in identifying citation functions. For example, *essential* in this example is not included to show the writer's *endorsement* or support of the source's proposition; Rather, it is used to contradict the source's claim (*difficult to describe* versus *essential*), thus to Engage with the cited view. The analyst's intervention therefore, is crucial in verifying the results of any automatic classifier of citation functions.

Of course, it is not easy to automate the classification of citation functions. However, as found in some classification schemes (e.g. Garzone and Mercer 2000), it is possible to a certain extent to produce preliminary lists of citations according to their functions. Once the citations are coded for form, stance and function, the analyst can perform a random sample check to verify the results of this coding. The use of such tools should also enable language researchers to observe move patterns in blocks and/or to search

for new stance and function combinations. For example, it is possible to identify new patterns when considering MF citations (e.g. *acknowledge*-to-Position-and-Engage). In this thesis, it was decided not to include such patterns due to time and space restrictions. However, it would be useful for corpus-based investigations to observe and compare such patterns to see how MF citations relate to preceding and succeeding texts.

Another issue with citation data is that it is not easy or even possible to produce a complete list of stance markers (e.g. reporting verbs, nouns or structures). Even if it were possible, it remains difficult to classify such markers into positive, negative or neutral markers. Therefore, from a purely linguistic point of view, the current thesis has made no attempt to automate the identification of blocks, citation functions and/or stance and function combinations. This is partly because the corpus was relatively small and it was of more interest to apply and revise the model categories based on analyses of the 1186 citations. Also, the current thesis uses a text-based approach rather than a corpus based approach. Another reason is that it was not possible to automate, or attempt to automate, the identification of blocks or the stance-function combinations until the complete model was developed, with the conventional and unconventional combinations, which was not the case until towards the end of the process. Nevertheless, if resources allowed (e.g. in CL or LIS), it would be useful to investigate citation practices in a larger corpus, which would probably require the allocation of tasks between a number of researchers. As noted above, one issue with large corpora is that it is difficult and sometimes impossible to investigate language in context (see Hunston 2002; Handford 2010 for more detail). The analyst will always play a vital role in verifying the results of such corpus-based investigations, even though their involvement is reduced with a larger corpus.

In LIS, there has been no agreement in the various classification schemes on the definition and/or identification of citation functions. Therefore, every classification scheme identifies citations according to the framework it adopts. In terms of the pedagogical implications of the current study for LIS research, one question is particularly important. The question relates to the extent to which the proposed model can be used to improve citation indices and other automatic retrieval systems. Of course, in order to improve h-indices in light of the proposed model, the model has to be completely automated. Therefore, the focus here is not to articulate how this can be achieved, as this is fundamentally an issue for LIS research; instead, I aim to draw

attention to how the findings of the proposed model may be used to inform decisions in LIS research on citation value and research impact.

As noted in Chapter 2 (Section 2.2), the focus in LIS research has been on improving information retrieval and/or quantitative measures of researcher contribution. In recent years, there has been an increasing awareness of the need to automate the classification of citation functions (e.g. Bertin et al. 2016). However, as noted by Bertin et al. (2016), the persistent need for linguistic intervention is still acknowledged. Bertin et al. (2016) use NLP techniques and n-grams to arrive at a way of classifying citation functions. Yet, they acknowledge the importance of the rhetorical structure in deciding importance and also the need for better implementation of automated annotation of citation contexts by more "linguistically motivated approaches" (ibid: 1431).

The identification of blocks demonstrates the importance of applying the model to LIS research. One of the main difficulties with citation data, particularly in LIS and CL research, is the lack of classification schemes that investigate citation practices beyond the sentence-level. As shown in various places in this thesis, this casts doubt on the classifications and their findings. In order to illustrate this, I draw on two examples from Chapter 4 (Example 4.37: *Waismann argued …. For this reason, …* and Example 4.39: *First estimates … were presented by REFs. These studies … share a lack of detail for …*.).

Any automated retrieval system would analyse the above citations at the sentence-level. Such systems would classify both examples as *acknowledging* external references (*Waismann* in 4.37 and *REFs* in 4.39) and therefore the assumed function might be to Signpost readers to external sources. Along these lines, examples such as Example 7.3 (and the second statement of Example 7.2) would be classified as simply *distancing* the source's views and therefore showing external Positions. Only the block-level analysis shows the writers' Engagement with the sources (e.g. criticising the references and showing them to be lacking in depth in Example 4.39) and/or their use of *acknowledged* views in Support of their arguments, as in Example 4.37. The proposed model thus provides an opportunity for LIS research to reconsider scientific contributions in light of the different roles performed by the different citations in the citing texts.

The extent to which research produced in this thesis can be thoroughly or partly automated is a secondary issue to the investigation and re-identification of citation role

in context. Overall, it could be argued that the identification of citation blocks can be partly automated. It is possible to teach computers to identify different citation instances and produce different blocks (whether each block consists of 3, 4 or even 5 paragraphs). At the moment, it is possible for some corpus tools to identify line and sentence breaks. It therefore seems possible for similar or future tools to mark paragraph endings and to group paragraphs for every citation in order to produce potential blocks.

Since it is argued that the text-internal functions are of more value than the text-external functions, it seems reasonable to argue that an *acknowledged*-to-Support citation is more important than an *endorsed*-to-Signpost citation. This is because the former combination, with Support as the function, plays a more prominent role in the build-up of the writer's argument and, more importantly, because it links the cited propositions to other elements within the citing text.

It should be noted, however, that the identification and classification of citation functions in terms of importance is not as clear-cut as one might have presumed. This is because the identification of the value of a pattern depends on a number of factors such as the citation being conventional or unconventional and the type of genre under investigation (e.g. empirical versus policy discussion research articles). One might argue that the function Engage is more important in policy discussion papers because it is natural for such articles to Engage with existing policies or government agendas (e.g. see Chapter 6). However, one important aspect in this regard relates to the kind of research environments being construed by the writers (cf. Beckerman and Pasek 1995 and McAlpine et al. 2009 in Chapter 6). In general, it seems reasonable to propose an interim hierarchy at this stage from the most to the least important functions: Support, Engage, Position and Signpost. However, this assumption needs to be tested on large corpora and different disciplines and genres before conclusions can be reached about citation functions and their value in different texts.

The current thesis emphasises the different values of the various research contributions in terms of citation loads. It argues that some citations are highly loaded (Support and Engage) for value while others are less loaded (Position) or even unloaded for value (Signpost) with respect to the writer's argument in the block. In order to demonstrate what is meant by this assertion, let us assume that there are two authors with two different citation indices. The citation index for the first author is 50 while the second

author has an index of 100. The first author is cited 40 times (80%) in Support of the citing works and 10 times (20%) to show the author's Position. The second author is cited 80 times (80%) to Signpost readers to their research, 10 times (10%) in Engagement with the writer's views and another 10 citations (10%) to show their Position. In this case, the first author is cited less than the second. However, the arguments of the first author more often play a crucial role of the citing texts whereas the second author is often cited to Signpost readers to their research. In this case, the citations of the first author would be highly (positively) loaded for value (80%) while the second author does not have a high load of citations (only 10%), negatively loaded even though they have a higher h-index.

The issue with this view is the extent to which it has the propensity to be automated and applied on a global and not individual scale to re-assess citations in terms of positive and negative loads. It would be premature at this stage, from a linguistic point of view, to assume that the model can solve the identified issues related to automatic identification. However, it suffices to argue that the proposed model demonstrates that citations differ in value according to the role they perform in the citing texts. Finding ways to automate the model would be promising for LIS research.

The ideal position for automatic retrieval systems or quantitative measures of scientific contribution would be to be able to identify citation loads for different authors. However, context (block) analysis is always needed to identify the citation functions and their roles in texts. The question that remains is whether we can automate the identification and analysis of citation functions in blocks. As noted in the implications for CL research, it seems reasonable to argue that block identification is achievable. The automatic identification and classification of citation functions depends on the extent to which corpus tools can code textual and interpersonal cues within and beyond the clause-level. Thompson et al.'s (2017) notions of explicit and implicit argumentation are of particular importance to the coding process as it is difficult on some occasions to identify the links drawn between clauses, propositions and paragraphs. However, the writer's stance usually becomes overt at some stage in argumentation and the block should include at least one marker of the writer's position in relation to the cited works. There are cases in which the writers make their argumentation explicit at the clause-level of every citation (e.g. Example 7.2). In such cases, the block-level choices would have to take into consideration those at the clause-

level for every citation to deduce the citations functions. For example, the clause-level choices of every citation in Example 7.2 gives an indication of the writers' overall position towards the cited works and their role in the citing text. In that example, the writers Engage with the *distanced* citations and this Engagement, unlike Example 7.3, is identified at the level of other citations. If no links can be identified between the citing and cited texts, then it seems reasonable to code such instances as examples of Signpost or Position depending on how the information is presented at the clause-level. For example, an *acknowledged* source functions to Signpost readers while an *acknowledged* proposition or argument would function to show the source's Position. The analyst can always verify the accuracy of the manual or automatic coding and provide labels for some doubtful cases (e.g. MF citations).

Notes

1- Though Moravcisk and Murugesan (1975) were the first to highlight the roles played by citations in the citing texts (see Chapter 2).
2- Students should also be exposed to implicit stances (see examples in the second half of Section 4.2). However, implicit stance can be studied and identified at the clause-level. For this reason, it has not been included in the block-level pedagogical implications in Sub-section 7.4.1.

Chapter Eight

Concluding Remarks: Summary and Recommendations for Further Research

In conclusion, the current thesis has adopted the constellation approach from a corpus study (Thompson et al. 2017) to compile a corpus of GEC papers and to propose a new model for investigating citation practices in interdisciplinary discourse. Thompson et al. establish that constellations can be found in GEC. Papers in each constellation share a particular disciplinary profile. I have selected constellation 1 as the most 'science-like' and constellation 5 as the most 'social science-like'. My own expectation was that the citation practices in the two constellations would differ qualitatively and quantitatively, and would reflect the "broader disciplinary communities from which they come" (Hunston 2018, personal communication).

In order to add a diachronic aspect to the analysis, I compared two time periods (1990-95 and 2008-10). As indicated in Chapter 3, I had no control, nor had Thompson et al. (2017) over the distribution of papers across the constellations or time periods. Therefore, I decided to sample the constellations and time periods evenly (for further explanation see Chapter 3; e.g. the early period of constellation 5 contains only five papers). Once the papers had been selected, another aim of the study was to observe synchronic and diachronic variation between the four data sets in terms of citation practices.

In order to observe the expected variation, a new model was proposed for the investigation of citation practices and their roles in texts. The proposed model builds on, extends and modifies to a certain extent some aspect of the adopted models on the analysis of citation form, stance and function (see Chapter 3). The ultimate goal is, as articulated elsewhere in this thesis, to provide a richer understanding of how citations act, and are acted upon, in the citing texts. To achieve this aim, a new unit for analysing citations was proposed: that is, the citation block. Simply put, it is argued that citations operate in blocks, rather than clauses, and for one to understand how they function in texts, they need to investigate their role at the block-level (see Chapter 3 for the identification of the block). That is not to imply, however, that clause-level investigations of citations are unimportant; rather, I would argue that the block-level

analysis complements, and on many occasions supersedes, the clause-level analysis of citations.

The proposed model can be seen as a triad system which combines form, stance and function analyses to better account for citation role in the citing texts. It is argued that this combination is useful for the understanding of how citations operate at different levels (clauses and blocks) and in different directions (i.e. how writers evaluate cited propositions and at the same time how the evaluated propositions function in the citing texts). When citing, academic writers draw on various resources within and beyond the clause-level to achieve one or more textual functions.

Combining the clause- and block-levels is highly important for the understanding of the role of citation in evaluation. The block-level analysis adds to our understanding on how attributed statements are evaluated beyond the immediate sentence-level. The previous research accounts for the ways in which attributed propositions are evaluated at the moment at which they are cited (e.g. via the use of evaluative reporting verbs). However, this thesis looks within and beyond the citing statements to observe the writer's signalling of their stance towards cited propositions at the block-level. The writer's stance towards the cited statements, which extends beyond the citing statement, is taken as an indication of the citation function. For example, if a writer *distances* a source at the clause-level and *contests* it at the block-level (see Example 7.3), then the negative stance at the block-level indicates that the writer is involved in critical dialogue with the citation, thus the function is Engage.

The block takes the emphasis beyond the clause-level choices (e.g. tense or reporting verbs) and focuses on the writer's stance and organisational divisions in the preceding and subsequent text to observe what effect the clause- and block-level choices have on the perception of the writer's overall position and the citation function(s). It should be noted in this regard that the identification and labelling of stance in this thesis has been limited to the clause-level in order to compare the clause- and block-level findings for every citation. This choice has resulted in identifying conventional and unconventional combinations, based on the analysis and coding of stance and function categories (see Sub-section 4.3 in Chapter 4). The conventional and unconventional combinations have various pedagogical implications. The block can be taken as the main unit of analysing citations when teaching reading, for example. EAP practitioners can design courses or

activities to raise students' awareness of how citations operate beyond the clause-level, not only to sensitise them to the available resources at each level, but also to teach them how knowledge is constructed at each level. This should gradually sensitise students to the effects of their choices and thus help them adapt to the practices of successful academic writing.

Identifying stance and function at the two levels is a complex process. In terms of stance, Hunston (2007: 39), building her view on Sinclair, points out that stance does not occur in "discrete items" and that it is a "cumulative" process. Also, Stubbs (1996: 92) contends that "words do not have determinate meanings which can be defined in advance of their use" and goes on to argue that words are flexible and "open to new uses" depending on the context in which they are used. Therefore, identifying stance (and the writer's position) in reporting can be highly subjective, even with the existence of certain linguistic cues at the clause-level of the citation.

The same argument applies to the identification of citation functions. Identifying the function(s) of a citation is a complex and multidimensional process which goes beyond the citer's intention/motive for citing and their stance at the clause-level. When combined, stance and function analyses reflect the role of citation (i.e. how a citation is evaluated and for what textual, not intended, purpose). Murakami et al. (2017), in their attempt to automate the identification of topics in GEC papers, argue that the identification of topics and topic transition is an "act of interpretation [that is] essentially qualitative" and which "should not be disguised by quantitative methods" (p. 273). Therefore, the authors call for manual labelling of topics in topic modelling. Similarly, I argue that the identification of citation functions has some level of subjectivity and cannot be completely automated. In other words, the analyst's intervention is likely to be required at some stage in the coding and labelling process. It has long been argued in LIS research (e.g. Garzone and Mercer 2000; Bornmann and Daniel 2008) that the identification of citation functions could be automated as long as the contextual markers are explicitly stated. However, I would argue that the different uses of the contextual markers and/or stance indicators might lead to inaccurate labelling of citation functions. This is not to say that automatic classifications are unfeasible, but I simply mean that human intervention will always be required. Therefore, automatic classifiers should take into consideration the fact that "the

meaning of a citation is highly context-bound and therefore generalizations about impact or influence are at best tentative" (Hellqvist 2009).

In order to emphasise the contextual nature of citations, I draw on Example 7.3 as a clear example in which the writer's argumentation is explicit. In 7.3, the writer uses clear markers of their stance and position at the clause- and block-levels (e.g. *claims, difficult to describe, however* and *essential*). It is unclear how automated classification schemes would classify examples such as 7.3, because of the contradictory signals of positive and negative evaluation (e.g. *difficult* and *essential*). The positive element (*essential*) is used to negatively evaluate the cited proposition. The default assumption in the literature is to treat this example as *distancing* the source's view, following the identification of the reporting verb *claims*.

The question remains open as to how function analysis can be automated in such examples or with examples which use adversative and additive cohesive devices (e.g. *Nevertheless, However* and *Indeed* as in Example 7.1). The argument in this chapter draws attention to one of Einstein's aphorisms, as reported in the analysis of corpus and discourse by Swales (2010: 293) and reiterated in Pecorari (2013: 156): "Sometimes that which counts cannot be counted and that which can be counted does not count". In Example 7.3, *essential* might be picked as a sign of positive evaluation of the first citation. However, the block-level analysis will show that it is included to contradict Smith's assertion in the first clause.

What should also be noted is that the writer makes explicit the links between clauses and his stance. Therefore, this is an example of explicit argumentation at the block-level. However, as shown in Chapters 4, 5 and 6, argumentation in citation blocks can also be implicit. Writers might not signal their position towards cited materials, allowing the reader to "infer" the writer's overall position towards the cited propositions (Du Bois 2007: 144). In such cases, the analyst's intervention is needed to account for citation function in the block.

Moving to the quantitative aspect of the study, the analysis has shown variation in citation practices between the papers in the early years and a convergence in the recent years. The variation can be attributed to the nature of the two constellations: constellation 1 represents the scientific component of GEC whereas constellation 5 represents the social scientific component. Knowledge construction in scientific writing

is deemed as value-free, impersonal and contingent on universally-accepted facts while in the social sciences it is usually personal, value-laden and contingent on argumentation (Becher and Trowler 2001). Stance is more explicit in the selected constellation 5 papers while it is implicit (tends to be realised at the surface of expansive stance types such as *acknowledge* or *distance*) in the constellation 1 papers. This finding could be attributed to the preference amongst the constellation 5 writers to use integral citations more than those of the constellation 1 papers. Integral forms facilitate the expression and identification of stance. The non-integral citations, however, are used more in the selected papers from constellation 1. It seems that this preference for non-integral forms has contributed to the increase in implicit stance in constellation 1 papers (see notes in Chapter 5) since such forms allow writers to give prominence to the cited research and concomitantly conceal their views towards the cited authors or their contribution.

The analysis also shows a diachronic change in the constellation 5 papers and convergence with the constellation 1 papers. This convergence suggests that GEC has been successful in developing its own voice or a more consistent style of citing, even though the evidence is only derived from the analysis of 1186 citations in twenty papers. It was originally expected that the constellation 5 papers will show more divergence from constellation 1 in the recent years because the journal explicitly states on its website that it welcomes contributions with a social scientific component. Therefore, the diachronic change was expected in the papers from the scientific component (constellation 1). The observed convergence in the recent years suggests that GEC is becoming more established as an interdisciplinary journal. In other words, the field was being established in the early years of the journal and thus the citation practices reflect variation between the papers from the two constellations. In recent years however, it could be argued that GEC became more established as an interdisciplinary journal and as a result, bridged the gap between citation practices in the two constellations.

This study provides small-scale evidence of variation between the two constellations. Analysis of a greater number of papers is required before reaching more robust quantitative conclusions about convergence or divergence between the constellations. The quantitative findings are intended to reflect synchronic and diachronic variation between the selected papers only. The discussions relating to the findings draw attention

to some potential reasons behind the observed differences in the citation practices of the two constellations.

The implications for pedagogy, CL and LIS could be considered as implications or recommendations for future research. The proposed model can be applied to a larger corpus, comprising other disciplines, genres or journals. Another possible direction is to examine move patterns in citation blocks across different disciplines (cf. Examples 6.1 and 6.4). Future studies may also apply certain aspects of the model, but differently. For example, stance analysis can be conducted at the clause- and block-levels of every citation to observe new patterns (e.g. *acknowledge*-to-*endorse*). Alternatively, both stance and function analyses can be conducted at the block-level. In such cases, examples such as 7.3 would be seen as examples of the *contest*-to-Engage, not *distance*-to-Engage, combination.

Stance analysis has been limited to the clause-level in this thesis to show that the analysis of each level can be misleading if taken in isolation of the accompanying level. The clause-level analysis of stance has the benefit of revealing variations which occur in the unconventional combinations (see Sub-section 4.3.2), between the identified stance and function categories. This should assist both students and teachers in the deconstruction of texts and the understanding of how language is constructed at each level.

There is a lack of diachronic (and also longitudinal) studies into citation practices. It would be useful for such studies to investigate diachronic change in citations within and across disciplines, subjects or genres. Such investigations might reveal convergence or divergence in citation practices and/or diachronic change in citation practices of a single discipline, journal or genre. Finally, if resources allowed, it would be extremely useful for research in LIS and/or CL to identify, and automate the identification of citations which are positively and negatively loaded for value. The text-internal functions Support and Engage have a high citation load, and are positively loaded for value. The text-external functions Position and Signpost are negatively loaded. The extent to which this is achievable depends on advancements in CL and LIS research.

References

Abbasi, AR, Akbari, N & Gravis, B 2006, 'Discourse appropriation, construction of identities, and the complex issue of plagiarism: ESL students writing in graduate school', *Journal of Second Language Writing*, vol. 15, pp. 102-117.

Adel, A 2014, 'Selecting quantitative data for qualitative analysis: a case study connecting a lexicogrammatical pattern to rhetorical moves', *Journal of English for Academic Purposes*, vol. 16, pp. 68-80.

Adel, A & Garretson, G 2006, Citation practices across the disciplines: the case of proficient student writing, in MC Pérez-Llantada Auria, RP Alastrue & CP Neumann (eds), *Academic and professional communication in the 21st century: genres, rhetoric and the construction of disciplinary knowledge*, proceedings of the 5th International AELFE Conference, Zaragoza.

Ahmed, T, Johnson, B, Oppenheim, C & Peck, C 2004, 'Highly cited old papers and the reasons why they continue to be cited: Part II. The 1953 Watson and Crick article on the structure of DNA', *Scientometrics*, vol. 61, no. 2, pp. 147-156.

Aikhenvald, AY 2004, *Evidentiality*, viewed 11 June 2018, Oxford University Press, Oxford, <https://research.jcu.edu.au/lcrc/storeroom/research-projects/evidentiality/folder-2-sashas-publications/evidentiality>.

Ali, NA 2016, 'Citation practices in the literature review: a contrastive study of Sudanese and British applied linguistics Ph.D. theses', *World Journal of English Language*, vol. 6, no. 4, pp. 25-49.

Aljabr, FS 2011, 'An analysis of cohesion and thematic development in university assignments written by undergraduate English major students in one Saudi university', Unpublished MA Thesis, University of Adelaide, Australia.

Amsterdamska, O & Leydesdorff, L 1989, 'Citations: indicators of significance', *Scientometrics*, vol. 15, no. 5-6, pp. 449-471.

Anderson, RC, Narin, F & McAllister, P 1978, 'Publication ratings versus peer ratings of universities', *Journal of the American Society for Information Science*, vol. 29, pp. 91-103.

Ard, J 1985, 'Vantage points for the analysis of scientific discourse', *English for Specific Purposes*, vol. 4, no. 3-20.

Aull, LL & Lancaster, Z 2014, 'Linguistic markers of stance in early and advanced academic writing a corpus-based comparison', *Written Communication*, vol. 31, pp. 151-183.

Baird, LM & Oppenheim, C 1994, 'Do citations matter?', *Journal of Information Science*, vol. 20, pp. 2-15.

Bakhtin, MM 1934-5/1981, *The dialogic imagination: four essays*, Edited by M. Holquist, transalted by C. Emerson & M. Holquist, University of Texas Press, Austin.

Bakhtin, MM 1986, *Speech genres and other late essays*, University of Texas Press, Austin.

Barton, E 1993, 'Evidentials, argumentation, and epistemological stance', *College English*, vol. 55, no. 7, pp. 745-769.

Barton, E 2002, Inductive discourse analysis: discovering rich features, in E Barton & G Stygall (eds), *Discourse studies in composition*, Hampton, pp. 19-42.

Basturkmen, H 2009, 'Commenting on results in published research articles and masters dissertations in language teaching', *Journal of English for Academic Purposes*, vol. 8, pp. 241-251.

Bazerman, C 1984, 'Modern evolution of the experimental report in Physics: Spectroscopic articles in Physical Review, 1983-1980', Social Studies of Science, vol. 14, no. 2, pp. 163-196.

Bazerman, C 1988, *Shaping written knowledge: the genre and activity of the experimental article in science*, The University of Wisconsin Press, US.

Becher, T & Trowler, PR 2001, *Academic tribes and territories: intellectual enquiry and the culture of disciplines*, 2nd edn, The Society for Research into Higher Education and Open University Press, England.

Beed, C & Beed, C 1996, 'Measuring the quality of academic journals: the case of economics', *Journal of Post Keynesian Economics*, vol. 18, no. 3, pp. 369-396.

Belter, CW 2014, 'Measuring the value of research data: a citation analysis of oceanographic data sets', *PLoS One*, vol. 9, no. 3, pp. 1-9.

Berkenkotter, C & Huckin, T 1995, *Genre knowledge in disciplinary communication: cognition/culture/power*, Erlbaum, Hillsdale, NJ.

Bertin, M, Atanassova, I, Sugimoto, CR & Lariviere, V 2016, 'The linguistic patterns and rhetorical structure of citation context: an approach using n-grams', *Scientometrics*, vol. 109, no. 3, pp. 1417-1434.

Biber, D 1988, *Variation across speech and writing*, Cambridge University Press, Cambridge.

Biber, D 1989, 'A typology of English texts', *Linguistics*, vol. 27, pp. 3-43.

Biber, D 2006, 'Stance in spoken and written university registers', *Journal of English for Academic Purposes*, vol. 5, no. 2, pp. 97-116.

Biber, D & Finegan, E 1988, 'Adverbial stance types in English', *Discourse Processes*, vol. 11, pp. 1-34.

Biber, D & Finegan, E 1989, 'Styles of stance in English: lexical and grammatical marking of evidentiality and effect', *Text*, vol. 9, no. 1, pp. 93-124.

Biber, D, Johansson, S, Leech, G, Conrad, S & Finegan, E 1999, *The Longman grammar of spoken and written English*, Longman, Harlow, UK.

Bloch, J 2010, 'A concordance-based study of the use of reporting verbs as rhetorical devices in academic papers', *Journal of Writing Research*, vol. 2, no. 2, pp. 219-244.

Bondi, M 2009, Historians at work: reporting frameworks in English and Italian book review articles, in K Hyland & G Diani (eds), *Academic evaluation: review genres in university settings*, Palgrave Macmillan, UK, pp. 179-198.

Bonzi, S & Snyder, HW 1991, 'Motivations for citation – A comparison of self citation and citation to others', Scientometrics, vol. 21, pp. 245-254.

Borg, E 2000, Citation practices in academic writing, in P Thompson (ed), *Patterns and perspectives: insights into EAP writing practice*, University of Reading, Reading, 27-45.

Bornmann, L & Daniel, H-D 2008, 'What do citation counts measure? A review of studies on citing behavior', *Journal of Documentation*, vol. 64, no. 1, pp. 45-80.

Broadus, RN 1971, 'Literature of social sciences – survey of citation studies', *International Social Science Journal*, vol. 23, pp. 236-243.

Brooks, TA 1984, 'Private acts and public objects: an investigation of citer motivation', *Journal of the America Society of Information Science*, vol. 36, pp. 223-229.

Brooks, TA 1985, 'Private acts and public objects: an investigation of the citer motivations', *Journal of the American Society for Information Science*, vol. 63, no. 4, pp. 223-229.

Brooks, TA 1986, 'Evidence of complex citer motivations', Journal of the American Society for Information Science, vol. 37, no. 1, pp. 34-36.

Brooks, TA 1988, 'Citer motivations', in A Kent (ed), *Encyclopaedia of library and Information Science*, vol. 43, pp. Dekker, New York, pp. 48-59.

Bryman, A 1988, *Quantity and quality in social research*, Unwin Hyman, London.

Brysse, K, Oreskes, N, O'Reilly, J & Oppenheimer, M 2013, 'Climate change prediction: erring on the side of least drama', *Global Environmental Change*, vol. 23, pp. 327-337.

Burrel, QL 2003, 'Predicting future citation behavior', *Journal of the American Society for Information Science and Technology*, vol. 54, pp. 372-378.

Butt, D, Fahey, R, Feez, S, Spinks, S & Yallop, C 2000, *Using functional grammar: an explorer's guide*, 2nd edn, MacMillan, Sydney.

Caldas-Coulthard, C.R. 1994, On reporting reporting: the representation of speech in factual and factional narratives, in M Coulthard (ed), *Advances in written text analysis*, Taylor and Francis Group, Routledge, London, 295-308.

Callon, M, Jean-Pierre, C, Turner, WA & Serge, B 1982, 'From translation to problematic networks: an introduction to co-word analysis, *Social Science Information*, vol. 22, pp. 191-235.

Campion, MA 1997, 'Rules for references: suggested guidelines for choosing literary citations for research articles in applied psychology', *Personnel Psychology*, vol. 50, no. 1, pp. 165-167.

Cano, V 1989, 'Citation behaviour: classification, utility, and location', *Journal of the American Society for Information Science*, vol. 40, pp. 284-290.

Case, DO & Higgins, GM 2000, 'How can we investigate citation behavior? A study of reasons for citing literature in communication', *Journal of the American Society for Information Science*, vol. 51, no. 7, pp. 635-645.

Chafe, W & Nichols, J 1991, 'Evidentiality: the linguistic coding of epistemology', *Language*, vol. 67, no. 1, pp. 133-141.

Chandrasoma, R, Thompson, C & Pennycook, A 2004, 'Beyond plagiarism: transgressive and non-transgressive intertextuality, *Journal of Language, Identity & Education*, vol. 3, no. 3, pp. 171-193.

Charles, M 2003, 'This mystery…': a corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines', *Journal of English for Academic Purposes*, vol. 2, no. 4, pp. 313-326.

Charles, M 2004, 'The construction of stance: a corpus-based investigation of two contrasting disciplines', Unpublished PhD thesis, University of Birmingham, UK.

Charles, M 2006, 'phraseological patterns in reporting clauses used in citation: a corpus-based study of theses in two disciplines', *English for Specific Purposes*, vol. 25, no. 3, pp. 310-331.

Charles, M 2007, 'Argument or evidence? Disciplinary variation in the use of the noun *that* pattern in stance construction', *English for Specific Purposes*, vol. 26, no. 2, pp. 203-218.

Chubin, DE & Moitra, SD 1975, 'Content analysis of references: adjunct or alternative to citation counting?', *Social Studies of Science*, vol. 5, pp. 423-441.

Coffin, C 2009, 'Incorporating and evaluating voices in a film studies thesis', *Writing & Pedagogy*, vol. 1, pp. 163-193.

Cole, S 2000, A short history of the use of citations as a measure of the impact of scientific and scholarly work, in B Cronin & HB Atkins (eds), *The web of knowledge. A festschrift in honor of Eugene Garfield*, Information Today, Medford, NJ.

Cole, S & Cole, JR 1967, 'Scientific output and recognition – study in operation of reward system in science', *American Sociological Review*, vol. 32, pp. 377-390.

Collins, HM 1985, 'The possibilities of science policies', *Social Studies of Science*, vol. 15, pp. 554-558.

Collins, HM 1999, 'Tantalus and the aliens: publications, audiences and the search for gravitational waves', *Social Studies of Science*, vol. 29, pp. 163-197.

Conrad, S & Biber, D 2000, Adverbial marking of stance in speech and writing, in S Hunston & G Thompson (eds), *Evaluation in text: authorial stance and the construction of discourse*, Oxford University Press, Oxford, pp. 56-73.

Coulthard, M 1994 (Ed.), *Advances in written text analysis*, Routledge, London.

Cozzens, SE 1989, 'What do citations count? The rhetorical-first model', *Scientometrics*, vol. 15, no. 5-6, pp. 437-447.

Cronin, B 1981, 'The need for a theory of citation', *Journal of Documentation*, vol. 37, no. 1, pp. 16-24.

Cronin, B 1982, 'Norms and functions in citation: the view of journal editors and referees in psychology', *Social Science Information Studies*, vol. 2, no. 2, pp. 65-78.

Cronin, B 1984, *The citation process: the role and significance of citations in scientific communication*, Taylor Graham, London.

Cronin, B 2005, *The hand of science: academic writing and its reward*, Scarecrow Press, Lanham, MD.

Crookes, G 1986, 'Towards a validated analysis of scientific text structure', *Applied Linguistics*, vol. 7, pp. 57-70.

Diani, G 2009, Reporting and evaluation in English book review articles: a cross-disciplinary study, in K Hyland & G Diani (eds), *Academic evaluation: review genres in university settings*, Palgrave Macmillan, UK, pp. 87-104.

Ding, Y, Liu, X, Guo, C & Cronin, B 2013, 'The distribution of references across texts: some implications for citation analysis', *Journal of Informetrics*, vol. 7, no. 3, pp. 583-592.

Ding, Y, Zhang, G, Chambers, T, Song, M, Wang, X & Zhai, C 2014, 'Content-based citation analysis: the next generation of citation analysis', *Journal of the Association for Information Science and Technology*, vol. 65, no. 9, pp. 1820-1833.

Dontcheva-Navratilova, O 2015, Cross-cultural variation in citation practices: a comparative analysis of citations in Czech English-medium and international English-medium linguistics journals, in R Plo Alastrué & C Pérez-Llantada (eds), *English as a scientific and research language: debates and discourses, English in Europe*, vol. 2, Mouton de Gruyter, Berlin, pp. 185-205.

Dontcheva-Navratilova, O 2016, 'Rhetorical functions of citations in linguistics research articles: a contrastive (English-Czech) study', Discourse and Interaction, vol. 9, no. 2, pp. 51-74.

Dressen-Hammouda, D 2014, 'Measuring the voice of disciplinarity in scientific writing: a longtitudinal exploration of experienced writers in geology', *English for Specific Purposes*, vol. 34, pp. 14-25.

Du Bois, JW 2007, The stance triangle, in R Englebretson (ed), *Stancetaking in discourse: subjectivity, evaluation, interaction*, John Benjamins Publishing, Amsterdam, 139-182.

Dudley-Evans, A 1986, Genre analysis: an investigation of the introduction and discussion sections of MSc. dissertations, in M Coulthard (ed), *Talking about text* (Discourse Monograph No. 13), English Language Research, University of Birmingham, 128-145.

Ean, JA 1982, Tense usage in reporting research in geotechnical writing, in *Working papers in ESL(2)*, University of Minnesota, Mineapolis, 72-91.

Eggins, S 2004, *An introduction to systemic functional linguistics*, 2nd edn, Continuum, London.

Englebretson, R 2007, Stancetaking in discourse: an introduction, in R Englebretson (ed), *Stancetaking in discourse: subjectivity, evaluation, interaction*, John Benjamins Publishing, Amsterdam, 1-25.

Errey, L 2007, What is it about other people's words?, in O Alexander (ed), *New approaches to materials development for language learning*, Peter Lang, Bern, 209-222.

Fairclough, N 1992, *Discourse and social change*, Polity Press, Cambridge.

Fardi, E, Tabatabaee Lotfi, S.A. 2017, 'The analysis of literature review sections of academic research articles by Iranian writers: citation analysis', *Modern Journal of Language Teaching Methods*, vol. 7, no. 2, pp. 107-116.

Flϕttum, K, Dahl, T & Kinn, T 2006, *Academic voices: across languages and disciplines*, vol. 148, John Benjamins, Amsterdam.

Flowerdew, L 2004, The argument for using English specialised corpora to understand academic and professional settings, in U Connor & T Upton (eds), *Discourse in the professions: perspectives from corpus linguistics*, John Benjamins, Amsterdam, pp. 11-36.

Forey, G, 'Aspects of theme and their role in workplace texts', Unpublished PhD thesis, University of Glasgow, UK.

Frost, CO 1979, 'The use of citations in literary research: a preliminary classification of citation functions', *The Library Quarterly: Information, Community, Policy*, vol. 49, no. 4, pp. 399-414.

Garfield, E 1962, 'Can citation indexing be automated?', *Essays of an Information Scientist*, vol. 1, pp. 84-90.

Garfield, E 1965, Can citation indexing be automated?, in ME Stevens, VE Giuliano & LB Heilprin (eds), *Statistical association methods for mechanical documentation, Symposium Proceedings*, National Bureau of Standards Miscellaneous Publication 269, National Bureau of Standards, Washington, DC, pp. 189-192.

Garfield, E 1972, 'Citation analysis as a tool in journal evaluation – journals can be ranked by frequency and impact of citations for science policy studies', *Science,* vol. 178, pp. 471-479.

Garfield, E 1998, 'Random thoughts on citationology: its theory and practice', *Scientometrics,* vol. 43, no. 1, pp. 96-76.

Garzone, M 1996, 'Automated classification of citations using linguistic semantic grammars', MSc Thesis, The University of Western Ontario, Canada.

Garzone, M & Mercer, RE 2000, 'Towards an automated citation classifier', Advances in Artificial Intellegence, *Conference of the Canadian Society for Computational Studies of Intelligence*, pp. 337-346.

Gil-Salom, L & Soler-Monreal, C 2014, 'Writers' positioning in literature reviews in English and Spanish computing doctoral theses', *Journal of English for Academic Purposes*, vol. 16, pp. 23-39.

Gilbert, GN 1977, 'Referencing as persuasion', *Social Studies of Science*, vol. 7, no. 1, pp. 113-122.

Global Environmental Change 2018, *Global Environmental Change: Human and Social Dimensions*, viewed 5 May 2018, <https://www.journals.elsevier.com/global-environmental-change>.

Gol, AK, Hezareh, BG & Soghondikolaei, EM 2014, 'A contrastive study of rhetorical functions of citation in Iranian and international ELT Scopus journals', *Linguistics and Literature Studies*, vol. 2, no. 6, pp. 155-165.

Groom, NW 2000, Attribution and averral revisited: three perspectives on manifest intertextuality in academic writing, in P Thompson (ed), *Patterns and perspectives: insights into EAP writing practice*, University of Reading, Reading, 14-25.

Groom, NW 2005, 'Pattern and meaning across genre and disciplines: an exploratory study', *Journal of English for Academic Purposes*, vol. 4, no. 3, pp. 257-277.

Groom, NW 2007, 'Phraseology and epistemology in humanities writing: a corpus-driven study', Unpublished PhD Thesis, University of Birmingham, UK.

Groom, NW 2009, Phraseology and Epistemology in academic book reviews: a corpus-driven analysis of two humanities disciplines, in K Hyland & G Diani (eds), *Academic evaluation: review genres in university settings*, Palgrave Macmillan, UK, pp. 122-139.

Gross, PLK & Gross, EM 1927, 'College libraries and chemical education', *Science*, vol. 66, pp. 385-389.

Gubrium, J & Holstein, J 2009, *Analysing narrative reality*, Sage, Thousand Oaks, CA.

Gunawardena, CN 1989, 'The present perfect in the rhetorical divisions of biology and biochemistry journal articles', *English for Specific Purposes*, vol. 8, no. 3, pp. 265-273.

Hagstrom, WO 1971, ''Inputs, outputs, and the prestige of university science departments', *Sociology of Education*, vol. 44, pp. 375-397.

Halliday, MAK 1994, *An introduction to functional grammar*, 2nd edn, Edward Arnold, London.

Halliday, MAK & Matthiessen, CMIM 2004, *An introduction to functional grammar*, 3rd edn, Edward Arnold, London.

Handford, M 2010, What can a corpus tell us about specialist genres?, in A O'Keeffe & M McCarthy (eds), *The Routledge handbook of corpus linguistics*, Routledge, London, pp. 255-269.

Hanney, S, Frame, I, Grant, J, Green, P & Buxton, M 2003, *From bench to bedside: tracing the payback forwards from basic or early clinical research – A preliminary exercise and proposals for a future study, HERG Research Report no. 31*, Health Economics Research Group, Brunel University, Uxbridge.

Harwood, N 2008a, 'Citers' use of citees' names: findings from a qualitative interview-based study', *Journal of the American Society for Information Science and Technology*, vol. 59, no. 6, pp. 1007-1011.

Harwood, N 2008b, 'Publication outlets and their effect on academic writers' citations', *Scientometrics*, vol. 77, no. 2, pp. 253-265.

Harwood, N 2009, 'An interview-based study of the functions of citations in academic writing across two disciplines', *Journal of Pragmatics*, vol. 41, pp. 497-518.

Harwood, N 2010, Research-based materials to demystify academic citation for postgraduates, in N Harwood (ed), *English language teaching materials: theory and practice*, Cambridge University Press, New York, 301-321.

Harwood, N & Petrić, B 2012, 'Performance in the citing behaviour of two student writers', *Written Communication*, vol. 29, no. 1, pp. 55-103.

Hellqvist, B 2009, 'Referencing in the humanities and its implications for citation analysis', *Journal of the American Society for Information Science and Technology*, vol. 61, no. 2, pp. 310-318.

Herlach, G 1978, 'Can retrieval of information from citation indexes be simplified? Multiple mentions of a reference as a characteristic of the link between cited and citing article', *Journal of the American Society for Information Science*, vol. 29, no. 6, pp. 308-310.

Hoey, M 2005, *Lexical priming: a new theory of words and language*, Routledge, London.

Holmes, J & Nesi, H 2009, Verbal and mental processes in academic disciplines, in M Charles, S Hunston & D Pecorari (eds), *Academic writing: at the interface of corpus and discourse*, Continuum, London, pp. 58-72.

Hood, S 2004, 'Appraising research: taking a stance in academic writing', Unpublished PhD thesis, University of Technology, Sydney.

Hooten, PA 1991, 'Frequency and functional use of cited documents in information science', *Journal of the American Society for Information Science*, vol. 42, no. 6, pp. 397-404.

Hu, G & Wang, G 2014, 'Disciplinary and ethnolinguistic influences on citation in research articles', *Journal of English for Academic Purposes*, vol. 14, pp. 14-28.

Hunston, S 1989, 'Evaluation in experimental research articles', Unpublished PhD thesis, University of Birmingham, UK.

Hunston, S 1993, Professional conflict: disagreement in academic discourse, in M Baker, G Francis & E Tognini-Bognelli (eds), *Text and technology: in honour of John Sinclair*, Benjamins, Amsterdam, pp. 115-134.

Hunston, S 1994, Evaluation and organization in a sample of written academic discourse, in M Coulthard (ed), *Advances in written text analysis*, Routledge, London, 191-218.

Hunston, S 1995, 'A corpus study of some English verbs of attribution', *Functions of Language*, vol. 2, no. 2, pp. 133-158.

Hunston, S 2000, Evaluation and the planes of discourse, status and value in persuasive texts, in S Hunston, & G Thompson (eds), *Evaluation in text: authorial stance and the construction of discourse*, OUP, Oxford, pp. 176-207.

Hunston, S 2002, *Corpora in applied linguistics*, Cambridge University Press, Cambridge.

Hunston, S 2004, 'It has rightly been pointed out …': attribution, consensus and conflict in academic discourse, in M Bondi, L Gavioli & M Silver, (eds), *Academic discourse, genre and small corpora*, Officina Edizioni, Roma, Italy, pp. 15-33.

Hunston, S 2007, Using a corpus to investigate stance quantitatively and qualitatively, in R Englebretson (ed), *Stancetaking in discourse: subjectivity, evaluation, interaction*, John Benjamins Publishing, Amsterdam, 27-48.

Hunston, S 2008, Collection strategies and design decisions, in A Ludeling & M Kyto (eds), *Corpus linguistics: an international handbook*, vol. 1, Walter de Gruyter, Berlin, pp. 154-167.

Hunston, S 2011, *Corpus approaches to evaluation: phraseology and evaluative language*, Routledge, New York.

Hunston, S & Murakami, A 2015, 'Multi-dimensional analysis, 'constellations', and interdisciplinary discourse', PowerPoint presentation, viewed 14 July 2015, University of Birmingham, UK.

Hunston, S & Thompson, G 2000 (eds), *Evaluation in text: authorial stance and the construction of discourse*, Oxford University Press, Oxford.

Hyland, K 1994, 'Hedging in academic writing and EAP textbooks', *English for Specific Purposes*, vol. 13, no. 3, pp. 239-256.

Hyland, K 1998, *Hedging in scientific research articles*, *John Benjamins Publishing*, Amsterdam.

Hyland, K 1999, 'Academic attribution: citation and the construction of disciplinary knowledge', *Applied Linguistics*, vol. 20, no. 3, pp. 341-367.

Hyland, K 2000, *Disciplinary discourses: social interactions in academic writing*, Pearson Education Limited, Longman, London.

Hyland, K 2001, 'Bringing in the reader: addressee features in academic articles', *Written Communication*, vol. 18, no. 4, pp. 549-574.

Hyland, K 2002, Activity and evaluation: reporting practices in academic writing, in J Flowerdew (ed), *Academic discourse*, Longman, London.

Hyland, K 2003, 'Self-citation and self-reference: credibility and promotion in academic publication', *Journal of the American Society for Information Science and Technology*, vol. 54, pp. 251-259.

Hyland, K 2004, Patterns of engagement: dialogic features and L2 undergraduate writing, in LJ Ravelli & RA Ellis (eds), *Analysing academic writing: contextualised frameworks*, Continuum, London, pp. 5-23.

Hyland, K 2005a, A convincing argument: corpus analysis and academic persuasion, in U Connor & T Upton (eds), *Discourse in the professions: perspectives from corpus linguistics*, John Benjamins, Amsterdam, pp. 87-114.

Hyland, K 2005b, 'Stance and engagement: a model of interaction in academic discourse', *Discourse Studies*, vol. 7, no. 2, pp. 173-192.

Hyland, K 2005c, *Metadiscourse*, Continuum, London, UK.

Hyland, K 2009, Corpus informed discourse analysis: the case of academic engagement, in M Charles, S.Hunston & D Pecorari (eds), *Academic writing: at the interface of corpus and discourse*, Continuum, London, pp. 110-128.

Hyland, K 2012, Undergraduate understandings: stance and voice in final year reports, in K Hyland & C.Sancho Guinda (eds), *Stance and voice in written academic genres*, Palgrave Macmillan, UK, pp. 134-150.

Hyland, K & Jiang, F.K 2016, 'Change of attitude? a diachronic study of stance', *Written Communication*, vol. 33, no. 3, pp. 251-274.

Hyland, K & Tse, P 2004, 'Metadiscourse in Academic writing: a reappraisal', *Applied Linguistics*, vol. 25, no. 2, pp. 156-177.

Jabbour, G 1997, 'Corpus linguistics, contextual collocation and ESP syllabus creation: a text-analysis approach to the study of medical research articles', Unpublished PhD thesis, University of Birmingham, UK.

Jacoby, S 1987, References to other researchers in literary research articles', *ELR Journal*, vol. 1, pp. 33-78.

Jafarigohar, M & Mohammadkani, A 2015, 'Reporting verbs in applied linguistics research articles by native and non-native writers', *Theory and Practice in Language Studies*, vol. 5, no. 12, pp. 2490-2496.

Jalilifar, A 2012, 'Academic attribution: citation analysis in master's theses and research articles in applied linguistics', *International Journal of Applied Linguistics*, vol. 22, no. 1, pp. 23-41.

Jalilifar, A & Dabbi, R 2012, 'Citation in applied linguistics: analysis of introduction sections of Iranian Master's theses', *Linguistik Online*, vol. 57, no. 7, pp. 91-104.

Johnstone, B 2009, Stance, style, and the linguistic individual, in A Jaffe (ed), *Stance: sociolinguistic perspectives*, Oxford University Press, New York, 29-52.

Jordan, R 1992, *Academic writing course*, Nelson, London.

Jullian, PM 2011, 'Appraising through someone else's words: the evaluative power of quotations in news reports', Discourse & Society, vol. 22, no. 6, pp. 766-780.

Kafes, H 2017, 'Citation practices among novice and expert academic writers', *Education and Science*, vol. 42, no. 192, pp. 441-462.

Kurtz, MJ, Eichhorn, G, Accomazzi, A, Grant, C, Demleitner, M, Murray, SS, Martimbeau, N & Elwell, B 2005, 'The bibliometric properties of article readership information', *Journal of the American Society for Information Science and Technology*, vol. 56, pp. 111-128.

Lackstrom, JE, Selinker, L & Trimble, LPG 1972, 'Grammar and technical English', English Teaching Forum, vol. 10, no. 5, (Part printed in J Swales (Ed.) 1985, *Episodes in English*, Pergamon, Oxford.

Lancaster, CI 2012, 'Stance and reader positioning in upper-level student writing in political theory and economics', Unpublished PhD thesis, The University of Michigan, USA.

Latour, B 1987, 'Science in action: how to follow scientists and engineers through society', Harvard University Press, Cambridge, MA.

Lawani, SM & Bayer, AE 1983, 'Validity in citation criteria for assessing the influence of scientific publications: new evidence with peer assessment', *Journal of the American Society for Information Science*, vol. 34, pp. 59-66.

Lee, DYW 2008, Corpora and discourse analysis: new ways of doing old things, in V Bhatia, J Flowerdew & H Jones (eds), *Advances in discourse studies*, Routledge, London, pp. 86-99.

Lee, SH 2010, 'Attribution in high- and low-graded persuasive essays', *Functions of Language*, vol. 17, pp. 181-206.

Leydesdorff, L 1998, 'Theories of citation?', *Scientometrics*, vol. 43, no. 1, pp. 5-25.

Lin, C, Chen, Y & Chang, C 2013, 'Citation functions in social sciences and humanities: preliminary results from a citation context analysis of Taiwan's history research journals', *Information Science and Technology*, vol. 50, no. 1, pp. 1-5.

Lipetz, BA 1965, 'Improvement of the selectivity of citation indexes to science literature through inclusion of citation relationship indicators', *Journal of the American Society for Information Science and Technology*, vol. 16, no. 2, pp. 81-90.

Liu, MX 1993a, 'Progress in documentation the complexities of citation practice: a review of citation studies', *Journal of Documentation*, vol. 49, no. 4, pp. 370-408.

Liu, MX 1993b, 'A study of citing motivation of Chinese scientists', *Journal of Information Science*, vol. 19, pp. 13-23.

Lopez, GS 1982, 'Article introductions in Spanish: a study in comparative rhetoric', Unpublished MA Thesis, Aston University, UK.

Luke, B & Kearings, K 2012, 'Attribution of words versus attribution of responsibilities: academic plagiagrism and university practice', *Organization*, vol. 19, no. 6, pp. 881-889.

Luukkonen, T 1990, 'Citations in the rhetorical, reward, and communication systems of science', Unpublished PhD Thesis, University of Tampere, Tampere.

Luukkonen, T 1997, 'Why has Latour's theory of citaitons been ignored by the bibliometric community? Discussion of sociological interpretations of citation analysis', *Scientometrics*, vol. 38, no. 1, pp. 27-37.

Lynn, FB 2014, 'Diffusing through disciplines: insiders, outsiders, and socially influenced citation behavior', *Social Forces*, vol. 93, no. 1, pp. 355-382.

McCain, KW & Turner, K 1989, 'Citation context analysis and aging patterns of journal articles in molecular genetics', *Scientometrics*, vol. 17, no.1-2, pp. 127-163.

McIntyre, D & Walker, B 2011, 'Discourse presentation in early modern English writing: a preliminary corpus-based investigation', *International Journal of Corpus Linguistics*, vol. 16, no. 1, pp. 101-130.

MacRoberts, MH & MacRoberts, BR 1984, 'The negational reference or the art of dissembling', *Social Studies of Science*, vol. 14, no. 1, pp. 91-94.

MacRoberts, MH & MacRoberts, BR 1989, 'Problems of citation analysis: a critical review', *Journal of the American society for Information Science*, vol. 40, no. 5, pp. 342-349.

MacRoberts, MH & MacRoberts, BR 1997, 'Citation analysis of a botany journal', *Journal of the American Society of Information Science*, vol. 48, pp. 274-275.

Maher, P 2015, 'The role of 'that' in managing averrals and attributions in post-graduate academic legel texts', *English for Specific Purposes*, vol. 40, pp. 42-56.

Maher, P 2016, 'The use of semi-technical vocabulary to understand the epistemology of a disciplinary field', *Journal of English for Academic Purposes*, vol. 22, pp. 92-108.

Mahlck, P & Persson, O 2000, 'Socio-bibliometric mapping of intra-departmental netweorks', *Scientometrics*, vol. 49, pp. 530-540.

Malcolm, L 1987, 'What rules govern tense usage in scientific articles', *English for Specific Purposes*, vol. 6, no. 1, pp. 31-43.

Malmström, H 2007, 'Accountability and the making of knowledge statements: a study of academic discourse', MA thesis, Lund University.

Manan, NA & Noor, NM 2015, 'The use of integral citations in Master's degree theses', *International Journal of Education and Research*, vol. 3, no. 7, pp. 233-246.

Mansourizadeh, K & Ahmad, U 2011, 'Citation practices among non-native expert and novice scientific writers', *Journal of English for Academic Purposes*, vol. 10, no. 3, pp. 152-161.

Maričić, S, Spaventi, J, Pavičić, L & Pifat-Mrzljak, G 1998, 'Citation context versus the frequency counts of citation histories', *Journal of the American Society for Information Science*, vol. 49, no. 6, pp. 530-540.

Martin, JR 1992, *English text: system and structure*, John Benjamins Publishing, Amsterdam.

Martin, JR 2000, Beyond exchange: appraisal systems in English, in S Hunston & G Thompson (eds), *Evaluation in text: authorial stance and the construction of discourse*, Oxford University Press, Oxford, pp. 142-175.

Martin, JR, Matthiessen, CMIM & Painter, C 2010, *Deploying functional grammar*, The commercial Press, Beijing.

Martin, JR & White, PRR 2005, *The language of evaluation: appraisal in English*, Palgrave Macmillan, Basingstoke, UK.

Mason, J 1996, *Qualitative researching*, Sage, London.

Mastura, N & Azlan, I 2013, 'Citation typologies and rhetorical functions of citations in in Master dissertations', *Journal of Creative Practices in Language Learning and Teaching*, vol. 1, no. 2, pp. 61-74.

Mercer, RE & Di Marco, C 2004, 'A design methodology for a biomedical literature indexing tool using the rhetoric of science', *Biolink 2004: linking biological literature, Ontologies and databases: tools for users*, Boston, MA.

Merton, RK (Ed.) 1973, The sociology of science: theoretical and empirical investigations, University of Chicago Press, Chicago, IL.

Moed, HF 2005, *Citation analysis in research evaluation*, Springer, Netherlands.

Molino, A 2010, 'Personal and impersonal authorial references: a contrastive study of English and Italian linguistics research articles', *Journal of English for Academic purposes*, vol. 9, no. 2, pp. 86-101.

Moravcsik, MJ 1973, 'Measures of scientific growth', *Research Policy*, vol. 2, no. 3, pp. 266-275.

Moravcsik, MJ & Murugesan, P 1975, 'Some results on the function and quality of citations', *Social Studies of Science*, vol. 5, no. 1, pp. 86-92.

Mosegaard, HMB 1998, *The functions of discourse particles*, John Benjamins, Amsterdam.

Muguiro, N 2015, 'Rhetorical functions of citations in interdisciplinary discourse: preliminary findings in the field of Educational Neurosciences', Unpublished PhD Module, University of Birmingham, UK.

Murakami, A, Thompson, P, Hunston, S & Vajn, D 2017 "What is this corpus about?': using topic modelling to explore a specialised corpus', *Corpora*, vol. 12, no. 2, pp. 243-277.

Murugesan, P & Moravcsik, MJ 1978, 'Variation of the nature of citation measures with journals and scientific specialities'. *Journal of the American Society for Information Science*, vol. 29, pp. 141–147.

Ngah, ZA & Sze, GS 1997, 'Characteristics of citations used by humanities researchers', *Malaysia Journal of Library and information science*, vol. 2, no. 2, pp. 19-36.

Nicolaisen, J 2002, 'The J-shaped distribution of citedness', *Journal of Documentation*, vol. 58, pp. 383-395.

Ochs, E 1989 (Ed.), 'The pragmatics of affect', Special Issue of Text, vol. 9, no. 3.

O'Connor, J 1980, 'Citing statements – recognition by computer and use to improve retrieval', *Proceedings of the American Society for Information Science*, vol. 17, pp. 177-179.

O'Connor, J 1982, 'Citing statements: computer recognition and use to improve retrieval', *Information Processing and Management*, vol. 18, no. 3, pp. 125-131.

O'Connor, J 1983, 'Biomedical citing statements: computer recognition and use to aid full-text retrieval', *Information Processing and Management*, vol. 19, no. 6, pp. 361-368.

Okamura, A 2007, 'Writer's intention behind the citation forms in academic texts', *The Economic Journal of Takasaki City University of Economics*, vol. 49, no. 3,4, pp. 57-70.

Okamura, A 2008a, 'Use of citation forms in academic texts by writers in L1 & L2 context', *The Economic Journal of Takasaki City University of Economics*, vol. 51, no. 1, pp. 29-44.

Okamura, A 2008b, 'Citation forms in scientific texts: similarities and differences in L1 and L2 professional writing', *Nordic Journal of English Studies*, vol. 7, no. 3, pp. 61-81.

Oppenheim, C 1995, 'The correlation between citation counts and the 1992 research assessment exercise ratings for British library and information science university departments', *Journal of Documentation*, vol. 51, pp. 18-27.

Oppenheim, C 1997, 'The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and archaeology', *Journal of Documentation*, vol. 53, pp. 477-487.

Oppenheim, C & Smith, R 2001, 'Student citation practices in an information science department', *Education for Information*, vol. 19, no. 4, pp. 299-323.

Oster, S 1981, 'The use of tenses in "Reporting Past Literature" in EST', in L Selinker, E Tarone and V Hanzeli (eds), *English for academic and technical purposes: studies in honor of Louis Trimble*, Newbury House, Rowley, pp. 76-90.

Peacock, M 2015, 'Stance adverbials in research writing', *lbérica*, vol. 29, pp. 35-62.

Pecorari, D 2003, 'Good and original: plagiarism and patchwriting in academic second-language writing, *Journal of Second Language Writing*, vol. 12, pp. 317-345.

Pecorari, D 2006, 'Visible and occluded citation features in postgraduate second-language writing', *English for Specific Purposes*, vol. 25, pp. 4-29.

Pecorari, D 2013, Additional reasons for the correlation of voice, tense and sentence function, in N-L Johannesson, G Melchers & B Björkman (eds), *Of butterflies and birds, of dialects and genres: essays in honour of Phillip Shaw*, Stockholm Studies in English (104), Stockholm University, pp. 153-167.

Pérez-Llantada, C 2009, Shifting identities, textual responses and conflicting demands in knowledge construction processes. In M. Gotti (Ed.), *Commonality and individuality in academic discourse*, Peter Lang, Frankfort, 191-213.

Petrić, B 2007, 'Rhetorical functions of citations in high- and low-rated master's theses', *Journal of English for Academic Purposes*, vol. 6, no. 3, pp. 238-253.

Petrić, B 2012, 'Legitimate textual borrowing: direct quotation in L2 students' writing', *Journal of Second Language Writing*, vol. 21, pp. 102-117.

Petrić, B & Harwood, N 2013, 'Task requirements, task representation, and self-reported citation functions: an exploratory study of a successful L2 student's writing', *Journal of English for Academic Purposes*, vol. 12, no. 2, pp. 110-124.

Pickard, V 1993, Should we be teaching refutation? Concordanced evidence from the field of applied linguistics, in N Bird, J Harris & M Ingham (eds), *Language and contest*, Institute of Language in Education, Hong Kong Education Department, Hong Kong, pp. 387-407.

Pickard, V 1995, 'Citing previous writers: what can we say instead of "say"?', *Hongkong Papers in Linguistics and Language Teaching*, vol. 18, pp. 89-102.

Pierce, SJ 1987, 'Characteristics of professional knowledge structures: some theoretical implications of citation studies', *Library and Information Science Research*, vol. 9, pp. 143-171.

Poppi, F 2004, Writer representation and authorial stance: a case study of the introductory chapters in economics textbooks, in M Bondi, L Gavioli & M Silver, (eds), *Academic discourse, genre and small corpora*, Officina Edizioni, Roma, Italy, pp. 127-137.

Pride, D & Knoth, P 2017a, 'Incidental or influential? Challenges in automatically detecting citation importance using publication full texts', Research and Advanced Technology for digital libraries, *International conference on Theory and Practice of Digital Libraries*, pp. 572-578.

Pride, D & Knoth, P 2017b, 'Incidental or influential? A decade of using text-mining for citation function classification', *16th International Society of Scientometrics and Informetrics Conference*, 16-20 Oct 2017, Wuhan.

Quirk, R, Greenbaum, S, Leech, G & Svartvik, J 1985, *A comprehensive grammar of the English language*, Longman, London.

Rabab'ah, G & Al-Marshadi, A 2013, 'Integrative vs non-integrative citations among native and non-native English writers', *International Education Studies*, vol. 6, no. 7, pp. 78-87.

Rabow, H 2005, 'The discovery of discoveries: exploring the dissemination of major findings in the life sciences', in P Ingwersen & B Larsen (eds), *Proceedings of the 10th International Conference of the International Society for Scientometrics and Informetrics*, Karolinska University Press, Stockholm.

Rego, ADMY 2001, 'Encapsulation and prospection in written scientific English', *English Studies at the Complutense University*, vol. 9, pp. 81-101.

Riddle, E 1986, 'The meaning and discourse function of the past tense in English', *TESOL Quarterly*, vol. 20, no. 2, pp. 267-286.

Riessman, CK 2008, *Narrative methods for the human sciences*, Sage, Thousand Oaks, CA.

Römer, U 2008, 'Identification impossible? A corpus approach to realisations of evaluative meaning in academic writing', *Functions of Language*, vol. 15, no. 1, pp. 115-130.

Samraj, B 2008, 'A discourse analysis of master's theses across disciplines with a focus on introductions', *Journal of English for Academic Purposes*, vol. 7, pp. 55-67.

Samraj, B 2013, 'Form and function of citations in discussion sections of master theses and research articles', *Journal of English for Academic Purposes*, vol. 12, no. 4, pp. 299-310.

Samraj, B & Monk, L 2008, 'The statement of purpose in graduate program applications: genre structure and disciplinary variation', *English for Specific Purposes*, vol. 27, no. 2, pp. 193-211.

Sandstorm, U, Wadskog, D & Karlsson, S 2005, Research institutes and universities: does collaboration pay?, in P Ingwersen & B Larsen (eds), *Proceedings of the 10th International Conference of the International Society for Scientometrics and Informetrics*, Karolinska University Press, Stockholm.

Schneider, JW 2006, 'Concept symbols revisited: naming clusters by parsing and filtering of noun phrases from citation contexts of concept symbols', *Scientometrics*, vol. 68, no. 3, pp. 573-593.

Scollon, R & Scollon, S 1997, 'Point of view and citation: fourteen Chinese and English versions of the 'same' news story', *Text*, vol. 17, no. 1, pp. 83-125.

Semino, E & Short, M 2004, *Corpus stylistics: speech, writing and thought presentation in a corpus of English writing*, Routledge, London.

Shadish, WR, Tolliver, D, Gray, M & Sen Gupta, SK 1995, 'Author judgement about works they cite: three studies from psychology journals', *Social Studies of Science*, vol. 25, no. 3, pp. 477-498.

Shahri, S & Soghondikolaei, M 2015, 'A comparative study of rhetorical functions of citation in applied linguistics research articles', *Proceedings of ADVED15 International Conference on Advances in Educaiton and Social Sciences*, 12-14 October, Istanbul, Turkey.

Shaw, P 1992, 'Reasons for the correlation of voice, tense, and sentence function in reporting verbs', *Applied Linguistics*, vol. 13, no. 3, pp. 302-319.

Shi, L 2008, 'Textual appropriation and citing behaviors of university undergraduates', *Applied Linguistics*, vol. 31, no. 1, pp. 1-24.

Shooshtari, ZG & Jalilifar, AR 2010, 'Citation patterns of research article discussions across subdisciplines of applied linguistics', *Journal of Teaching Language Skills*, vol. 2, no. 1, pp. 45-66.

Short, M & Semino, E & Wynne, M 2002, 'Revisiting the notion of faithfulness in discourse presentation using a corpus approach', *Language and Literature*, vol. 11, no. 4, pp. 325-355.

Silverman, D 2010, *Doing qualitative research: a practical handbook*, 3rd edn, Sage Publications Ltd., London.

Sinclair, JMcH. 1981, Planes of discourse. In S.N.A. Rizvi (Ed.), *The two-fold voice, essays in honour of Ramesh Mohan*, University of Salzburg, Salzburg, 70-89.

Sinclair, JMcH. 1985, On the integration of linguistic description, in T Van Dijk (ed), *Handbook of discourse analysis*, Academic Press, London, 13-28.

Sinclair, JMcH 1986a, Fictional worlds, in M Coulthard (ed), *Talking about text: studies presented to David Brazil on his retirement*, Discourse analysis monographs, no. 13, English Language Research, University of Birmingham, UK, 43-60.

Sinclair, JMcH 1986b, First throw away your evidence, in G Leitner (ed), *The English reference grammar*, Max Niemeyer Verlag, Tubingen.

Sinclair, JMcH 1987, *Mirror for a text*, MS. University of Birmingham, UK, 1-20.

Sinclair, JMcH 1988, 'Mirror for a text', *Journal of English and Foreign Languages*, vol. 1, University of Hyderabad, India, pp. 15-44.

Sinclair, JMcH 1991, *Corpus, concordance, Collocations*, Oxford University Press, Oxford.

Sinclair, JMcH 2004, *Trust the text: language, corpus and discourse*, Routledge, London.

Small, H 1982, Citation context analysis, in B Dervin & MJ Voight (eds), *Progress in communication sciences*, vol. III, Greenwood Publishing Group, Norwood, NJ, USA, pp. 287-310.

Smith, L 1981, 'Citation analysis', *Library Trends*, vol. 30, pp. 83-106.

Snodgrass, EA 2011, 'Referencing sources in discussion and conclusion sections', Unpublished MA thesis, Iowa State University, USA.

Soler-Monreal, C & Gil-Salom, L 2011, 'A cross-language study on citation practice in PhD theses', *International Journal of English studies*, vol. 11, no. 2, pp. 53-75.

Stubbs, M 1996, *Text and corpus analysis*, Blackwell Publishers Ltd., Oxford.

Sula, CA & Miller, M 2014, 'Citations, contexts, and humanistic discourse: toward automatic extraction and classification', *Literary and Linguistic Computing*, vol. 29, no. 3, pp. 425-464.

Swales, JM 1981, *Aspects of article introductions*, Language Studies Unit, University of Aston, Birmingham.

Swales, JM 1986, 'Citation analysis and discourse analysis', *Applied Linguistics*, vol. 7, no. 1, pp. 39-56.

Swales, JM 1990, *Genre analysis: English in academic and research settings*, Cambridge University Press, UK.

Swales, JM 2010, Afterword, in M Charles, D Pecorari & S Hunston (eds), *Academic writing: at the interface of corpus and discourse*, Continuum, London, pp. 291-294.

Swales, JM & Feak, C 2004, *Academic writing for graduate students*, University of Michigan Press, Ann Arbor.

Swales, JM & Najjar, H 1987, 'The writing of research article introductions', *Written Communication*, vol. 4, pp. 175-192.

Tadros, A 1985, *Prediction in text*, English Language Research, University of Birmingham, Birmingham, UK.

Tadros, A 1993, The pragmatics of text averral and attribution in academic texts. In M. Hoey (Ed.), *Data description discourse*, HarperCollins, London, 98-114.

Tafaroji Yeganeh, M & Boghayeri, M 2014, 'The frequency and function of reporting verbs in research articles written by native Persian and English speakers', paper presented at second global conference of linguistic foreign language teaching, Dubai, 11-13 December, retrieved from <https://ac.els-cdn.com/S1877042815035685/1-s2.0-S1877042815035685-main.pdf?_tid=003b02c0-9ab8-4f29-9ccf-cbbafea67083&acdnat=1535750058_47e721c297b4f68c54e85653dd507adf>.

Teufel, S, Siddharthan, A & Tidhar, D 2006, Automatic classification of citation function, in *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics*, Sydney, Australia, pp. 103-110.

Thompson, G 1994, *Reporting: helping learners with real English*, HarperCollins Publishers, London.

Thompson, G 1996, 'Voices in the text: discourse perspectives on language reports', *Applied Linguistics*, vol. 17, no. 4, pp. 501-530.

Thompson, G & Hunston, S 2000, Evaluation: an introduction, in S Hunston & G Thompson (eds), *Evaluation in text: authorial stance and the construction of discourse*, Oxford University Press, Oxford, pp. 1-27.

Thompson, G & Ye, Y 1991, 'Evaluation in the reporting verbs used in academic papers', *Applied Linguistics*, vol. 12, no. 4, pp. 365-382.

Thompson, G & Zhou, J 2000, Evaluation and organization in text: the structuring role of evaluative disjuncts, in S Hunston & G Thompson (eds), *Evaluation in text: authorial stance and the construction of discourse*, Oxford University Press, Oxford, pp. 121-141.

Thompson, P 2000, Citation practices in PhD theses, in L Burnard & T McEnry (eds), *Rethinking language pedagogy from a corpus perspective: papers from the Third International Conference on Teaching and Language Corpora, Lodz Studies in Language*, Peter Lang, Hamburg, pp. 91-102.

Thompson, P 2001, 'A pedagogically-motivated corpus-based examination of PhD theses: macrostructures, citation practices and uses of modal verbs', Unpublished PhD thesis, University of Reading, UK.

Thompson, P 2005a, 'Points of focus and position: intertextual reference in PhD theses', *Journal of English for academic purposes*, vol. 4, no. 4, pp. 307-323.

Thompson, P 2005b, Aspects of identification and position in intertextual reference in PhD theses, in E Tognini-Bonelli, & G Del Lungo Camiciotti (eds), *Strategies in academic discourse*, John Benjamins Publishing, Amsterdam, pp. 31-50.

Thompson, P 2009, Literature reviews in applied theses: evidence and problems, in K Hyland & G Diani (eds), *Academic evaluation: review genres in university settings*, Palgrave Macmillan, Basingstoke, UK, pp. 50-67.

Thompson, P 2012, Achieving a voice of authority in PhD theses, in K Hyland & C Sancho Guinda (eds), *Stance and voice in written academic genres*, Palgrave Macmillan, UK, pp. 119-133.

Thompson, P, Hunston, S, Murakami, A & Vajn, D 2017, 'Multi-dimensional analysis, text constellations, and interdisciplinary discourse', *International Journal of Corpus Linguistics*, vol. 22, no. 2, pp. 153-186 (pp. 1-34).

Thompson, P & Tribble, C 2001, 'Looking at citations: using corpora in English for academic purposes', *Language Learning & Technology*, vol. 5, no. 3, pp. 91-105.

Van Raan, AFJ 1988 (Ed.), *Handbook of quantitative studies of science and technology*, Elsevier, Amsterdam.

Van Raan, AFJ 2004, Measuring science. Capita selecta of current main issues, in HF Moed, W Glanzel & U Schmoch (eds), *Handbook of quantitative science and technology research: The use of publication and patent statistics in studies of S&T systems*, Kluwer Academic Publishers, Dordrecht.

Van Raan, AFJ, Visser, MS, Van Leeuwen, TN & Van Wijk, E 2003, 'Bibliometric analysis of psychotherapy research: performance assessment and position in the journal landscape', *Psychotherapy Research*, vol. 13, pp. 511-528.

Vinkler, P 1987, 'A quasi-quantitative citation model', *Scientometrics*, vol. 12, pp. 47–72.

Voloshinov, VN 1995, *Marxism and the philosophy of language, Bhakhtinian Thought – an introductory reader*. S. Dentith, L. Matejka & IR. Titunik (trans.), Routledge, London.

Voos, H & Dagaev, KS 1976, 'Are all citations equal? Or, did we op. cit. your idem?', *The Journal of Academic Librarianship*, vol. 1, no. 6, pp. 19-21.

Wan, X & Liu, F 2014, 'WL-Index: leveraging citation mention number to quantify an individual's scientific impact', *Journal of the Association for Information Science and Technology*, vol. 65, no. 12, pp. 2509-5217.

Wang, P & White, MD 1999, 'A cognitive model of document use during a research project, Study II, decisions at the reading and citing stages', *Journal of the American Society for Information Science*, vol. 50, no. 2, pp. 98-114.

Weissberg, R & Buker, S 1990, *Writing up research: experimental research report writing for students of English*, Prentice Hall Regents, Englewoods Cliff, NJ.

White, HD 2004, 'Citation analysis and discourse analysis revisited', *Applied Linguistics*, vol. 25, no. 1, pp. 89-116.

White, MD & Wang, P 1997, 'A qualitative study of citing behavior: contributions, criteria, and metalevel documentation concerns', *The Library Quarterly: Information, Community, Policy*, vol. 67, no. 2, pp. 122-154.

White, PRR 2002, 'Appraisal: the language of attitudinal evaluation and intersubjective stance', Appraisal Website, <www.grammatics.com/appraisal>.

White, PRR 2003, 'Beyond modality and hedging: a dialogic view of the language of intersubjective stance', *Text*, Special Edition on Appraisal, vol. 23, no. 3, pp. 259-284.

Widdowson, H 1998, 'Context, community and authentic language', TESOL Quarterly, vol. 32, no. 4, pp. 705-716.

Widdowson, H 2000, 'On the limitations of linguistics applied', Applied Linguistics, vol. 21, no. 1, pp. 3-25.

Williams, H 2010, 'Implicit attribution', *Journal of Pragmatics*, vol. 42, no. 3, pp. 617-636.

Woodward-Kron, R 2002, 'Critical analysis versus description: examining the relationship in successful student writing', *Journal of English for Academic Purposes*, vol. 1, no. 2, pp. 121-143.

Woodward-Kron, R 2003, 'Critical analysis and the journal article review assignment', *Prospect*, vol. 18, no. 2, pp. 20-36.

Woolgar, S 1991, 'Beyond the citation debate: towards a sociology of measurement technologies and their use in science policies', *Science and Public Policy*, vol. 18, pp. 319-326.

Zhang, G, Ding, Y & Milojević, S 2013, 'Citation content analysis (CCA): a framework for syntactic and semantic analysis of citation content', *Journal of the American Society for Information Science and Technology*, vol. 64, no. 7, pp. 1490-1503.

Zhao, D & Strotman, A 2014, 'In-text author citation analysis: feasibility, benefits, and limitations', *Journal of the Association for Information Science and Technology*, vol. 65, no. 11, pp. 2348-2358.

Zhao, SX, Tan, AM & Ye, FY 2012, 'Distributive h-indices for measuring multilevel impact', *Journal of the American Society for Information Science and Technology*, vol. 63, no. 10, pp. 2074-2086.

Zhu, X, Turney, P, Lemire, D & Vellino, A 2015, 'Measuring academic influence: not all citations are equal', *Journal of the Association of Information Science and Technology*, vol. 66, no. 2, pp. 408-427.

List of papers used in this thesis

Constellation 1 (1990-95)

Brookfield, H & Byron, Y 1990, 'Deforestation and timber extraction in Borneo and the Malay Peninsula: the record since 1965', *Global Environmental Change*, vol. 1, no. 1, pp. 42-56.

Millette, TM, Tuladhar, AR, Kasperson,, RE & Turner II, BL 1995, 'The use and limits of remote sensing for analysing environmental and social change in the Himalayan Middle Mountains of Nepal', *Global Environmental Change*, vol. 5, no. 4, pp. 367-380.

Pernetta, JC 1992, 'Impacts of climate change and sea-level rise on small island states: national and international responses', *Global Environmental Change*, vol. 2, no. 1, pp. 19-31.

Schulze, RE, Kiker, GA & Kunz, RP 1993, 'Global climate change and agricultural productivity in Southern Africa', *Global Environmental Change*, vol. 3, no. 4, pp. 330-349.

Smith, NJH, Alvim, P, Homma, A, Falesi, I & Serrão, A 1991, 'Environmental impacts of resource exploitation in Amazonia', *Global Environmental Change*, vol. 1, no. 4, pp. 313-320.

Constellation 1 (2008-10)

Dull, RA 2008, 'Unpacking El Salvador's ecological predicament: theoretical templates and "long-view" ecologies', *Global Environmental Change*, vol. 18, no. 2, pp. 319-329.

McAlpine, CA, Etter, A, Fearnside, PM, Seabrook, L & Laurance, WF 2009, 'Increasing world consumption of beef as a driver of regional and global change: a call for policy action based on evidence from Queensland (Australia)', Colombia and Brazil', *Global Environmental Change*, vol. 19, no. 1, pp. 21-33.

Phelps, J, Guerrero, MC, Dalabajan, DA, Young, B & Webb, EL 2010, 'What makes a 'REDD' country?', *Global Environmental Change*, vol. 20, no. 2, pp. 322-332.

Popp, A, Lotze-Campen, H & Bodirsky, B 2010, 'Food consumption, diet shifts and associated non-CO_2 greenhouse gases from agricultural production', *Global Environmental Change*, vol. 20, no. 3, pp. 451-462.

Schandl, H & West, J 2010, 'Resource use and resource efficiency in the Asia-Pacific region', *Global Environmental Change*, vol. 20, no. 4, pp. 636-647.

Constellation 5 (1990-95)

Beckerman, W & Pasek, J 1995, 'The equitable international allocation of tradable carbon emission permits', *Global Environmental Change*, vol. 5, no. 5, pp. 405-413.

Döös, BR 1994, 'Why is environmental protection so slow?', *Global Environmental Change*, vol. 4, no. 3, pp. 179-184.

Herrick, C & Jamieson, D 1995, 'The social construction of acid rain: some implications for science/policy assessment', *Global Environmental Change*, vol. 5, no. 2, pp. 105-112.

Kempton, W 1991, 'Lay perspectives on global climate change', *Global Environmental Change*, vol. 1, no. 3, pp. 183-208.

Shackley, S & Skodvin, T 1995, ‘IPCC gazing and the interpretative social sciences: a comment on Sonja Boehmer-Christiansen’s: ‘Global climate protection policy: the limits of scientific advice’’, *Global Environmental Change,* vol. 5, no. 3, pp. 175-180.

Constellation 5 (2008-10)

Grasso, M 2010, ‘An ethical approach to climate adaptation finance’, *Global Environmental Change,* vol. 20, no. 1, pp. 74-81.

Hagerman, S, Dowlatabadi, H, Satterfield, T & McDaniels, T 2010, ‘Expert views on biodiversity conservation in an era of climate change”, *Global Environmental Change,* vol. 20, no. 1, pp. 192-207.

Lahsen, M 2008, ‘Experiences of modernity in the greenhouse: a cultural analysis of a physicist “trio” supporting the backlash against global warming’, *Global Environmental Change,* vol. 18, no. 1, pp. 204-219.

Mortreux, C & Barnett, J 2009, ‘Climate change, migration and adaptation in Funafuti, Tuvalu’, *Global Environmental Change,* vol. 19, no. 1, pp. 105-112.

Räthzel, N & Uzzell, D 2009, ‘Changing relations in global environmental change’, *Global Environmental Change,* vol. 19, no. 3, pp. 326-335.