國立政治大學英國語文學系

博士學位論文

學術寫作中名詞補語結構的立場名詞後置修飾語：基於語料庫的跨學科和跨文化研究

Post-modification of stance nouns in *Noun Complement* construction: A corpus-based study of academic writing across disciplines and cultures

指導教授：劉怡君 博士

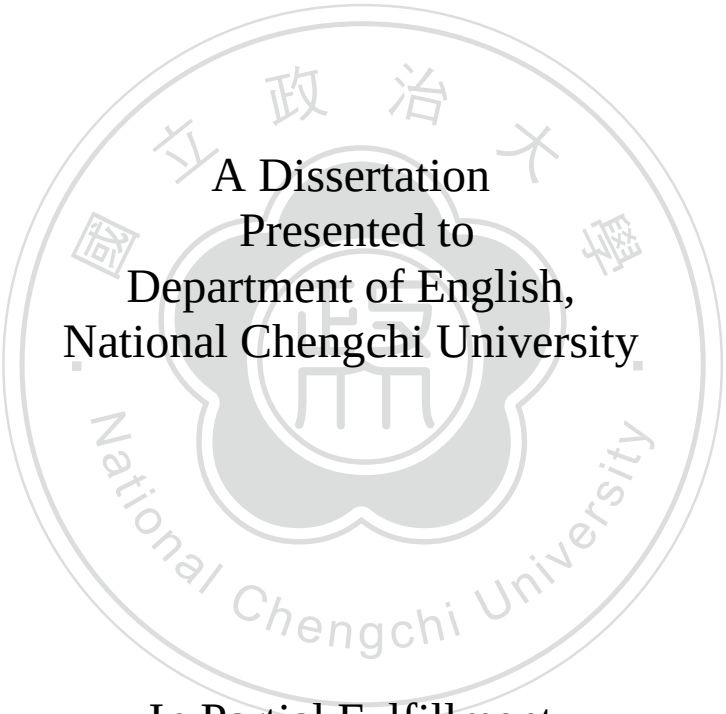
游曉曄 博士

研究生：黃廷 撰

中華民國 110 年 1 月

DOI:10.6814/NCCU202100169

Post-modification of stance nouns in *Noun Complement* construction: A corpus-based study of academic writing across disciplines and cultures



A Dissertation
Presented to
Department of English,
National Chengchi University

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

by

Ting Huang

January, 2021

DOI:10.6814/NCCU202100169

To My Dad

DOI:10.6814/NCCU202100169

ACKNOWLEDGEMENTS

First of all, I would like to thank my two supervisors, Professor Yichun Liu and Professor Xiaoye You, who have patiently guided me through the years of my doctoral study and inspired confidence in me. Both of them are diligent and professional scholars. Although they are always busy with teaching and doing research, they would still spare time discussing with me about my dissertation in the office or via Skype, whenever I made appointments with them. As professors, their kindness and gentleness have led me to take them as role models in my future life-long career. Without their guidance, I would never be able to finish my dissertation and get the degree. I owe them a great debt of gratitude!

I also want to express my gratitude to my parents and all my family. Without their support and encouragement, I might not be able to have such an opportunity to continue my study in the doctoral program. I love all my family.

Finally, I would thank all the professors, faculty members, classmates, and friends whom I have met and communed with on the campus or in the TESOL academic community, for your friendship and sincere support!

DOI:10.6814/NCCU202100169

TABLE OF CONTENTS

DOI:10.6814/NCCU202100169

DOI:10.6814/NCCU202100169

DOI:10.6814/NCCU202100169

DOI:10.6814/NCCU202100169

LIST OF TABLES

DOI:10.6814/NCCU202100169

DOI:10.6814/NCCU202100169

DOI:10.6814/NCCU202100169

LIST OF FIGURES

DOI:10.6814/NCCU202100169

DOI:10.6814/NCCU202100169

國立政治大學英國語文學系博士班

博士論文提要

論文名稱：學術寫作中名詞補語結構的立場名詞後置修飾語：基於語料庫的跨學科和跨文化研究

指導教授：劉怡君博士、游曉曄博士

研究生：黃廷

論文提要內容：

學術寫作不僅是事實的陳述更是作者透過研究來表達個人觀點與立場。因此，作者如何表達立場才能說服讀者認同其主張，已成為學術寫作研究中的重要議題。許多研究已探討了在構建立場時語言使用的特徵，如模糊限制語、報導動詞、指示詞、時態等。但是，名詞補語結構中立場名詞的後置修飾語卻被相對忽視了。本研究採用了跨國跨領域語料庫和混合研究方法來考察這種結構的使用。本研究的目標是從以下三個方面闡釋目標結構在不同學科和文化背景下使用的差異：（1）其詞彙語法特徵，（2）其功能，以及（3）其使用的潛在動機。

本研究自建語料庫包含 600 篇學術文章發表於英美和台灣與中國大陸的頂級期刊。為了跨領域對比後置修飾立場名詞補語，本資料庫所收集的學術文章分別來自於兩個學門：自然科學（化學、物理、生物）和社會科學（應用語言學、法律、經濟學）。透過推論統計方法，來確定不同學科領域和語言文化背景的學術作者在使用目標結構時的詞彙語法差異。本研究也採用定性分析的「功能分析架構」，以考察這些詞彙語法特徵發揮的功能在跨學科和跨文化方面的差異。此外，利用「潛

DOI:10.6814/NCCU202100169

在動機解釋模型」，探討不同學科領域和語言文化背景的作者使用該結構時的可能原因。研究結果表明，特定學科領域和民族文化在語言、修辭、策略和社會文化的實踐和結構存在差異。

本研究透過跨文化、跨學科的語料對比，探討名詞補語結構的後置修飾語使用差異。本研究自建的語料庫方法與分析模型對於目前英語學術寫作在語料庫立場分析研究和跨文化跨領域言詞分析研究上有實質重要貢獻。

關鍵字：後置修飾語，名詞補語結構，跨學科差異，跨文化差異，語料庫研究

DOI:10.6814/NCCU202100169

ABSTRACT

Academic writing is not simply concerned with the presentation of facts. In fact, academic writers tend to incorporate their own attitudes and judgements into the text. Thus, how to project an authorial stance that persuades readers of the writer's claim has become a pivotal element in academic written discourse. Different linguistic features, such as hedges, reporting verbs, directives, tense, have been examined for the roles they play in stance-making practices. A relatively neglected means of stance expression, however, has been the post-modification of the stance noun in the *Noun Complement* construction (as *common to many cell marking techniques* in *This procedure has the disadvantage common to many cell marking techniques that the cells selected for labelling need to be highly colchicine-resistant*). This study employed a corpus-based and mixed-methods approach to examining the use of this structure. The three major goals of this study are to describe how the target structure is used across different disciplinary and cultural contexts in terms of (1) its lexico-grammatical features, (2) its functions, and (3) the underlying motivations behind its use.

The research in this study is based on a self-built corpus of 600 samples of published research articles from the journals based in the UK or US and in Taiwan or mainland China in two branches of the natural (chemistry, physics, biology) and social sciences (applied linguistics, law, economics). The inferential statistical methods were applied to identifying the variances in the lexico-grammatical features of the target structure between disciplinary fields and academic writers of different native languages and cultures (English and Chinese). For the qualitative analyses, an analytical framework for description of the post-modification functions in *Noun Complement* construction was developed to analyze the cross-disciplinary and cross-cultural differences in the functions served by these statistically significant lexico-grammatical features. An explanatory model of the underlying motivations behind different uses of the target structure was also constructed to reveal the reasons that underlie the ways in which writers from contrasting disciplinary fields and linguistic-cultural backgrounds use this structure differently. Results show that the variations in disciplinary and cultural uses of post-modification can be attributed to the various factors in the linguistic, rhetorical, relational, and socio-cultural dimensions of the practices and structures of the particular disciplinary fields and cultures.

DOI:10.6814/NCCU202100169

In summary, this dissertation contributes to a better understanding of the *Noun Complement* construction in terms of its post-modification use from the cross-disciplinary and cross-cultural perspectives between the academic communities of the natural and social sciences in the native English countries and the Great China region. It also provides insight into pedagogical practice for new members of the Chinese disciplinary communities and into future research on the cross-disciplinary and cross-cultural use of *Noun Complement* construction in English academic writing.

Keywords: Post-modification, *Noun Complement* Construction, Cross-disciplinary Differences, Cross-cultural Differences, Corpus-based Study

DOI:10.6814/NCCU202100169

CHAPTER 1

Introduction

1.1 The need for studying post-modification in *Noun Complement* construction in academic writing

Academic writing offers scholars a stage to represent what the world is like (Hyland, 2000). However, writing in this form is not just about conveying facts but rather a "persuasive endeavour" (Jiang & Hyland, 2015, p. 529), seeking to lead the reader to the writer's opinion (e.g., Charles, 2006; Gilbert & Mulkay, 1984; Jiang, 2017; Latour, 1987; Myers, 1990). Following this generally accepted understanding of academic writing, researchers have shown an increasing interest in how academic writers incorporate their own "personal feelings, attitudes, value judgments, or assessments" (Biber et al., 1999: p. 966) into texts. A wide range of studies have been conducted in this line of research using various constructs, which include "stance" (e.g., Biber & Finegan, 1989; Charles, 2007; Jiang & Hyland, 2015), "evaluation" (e.g., Hunston, 1989, 1993, 1994; Theta, 1997), "appraisal" (e.g., Thompson & Hunston, 2000; White, 2002) and "metadiscourse" (Hyland, 2005b; Hyland & Guinda, 2012; Hyland & Jiang, 2018). Different linguistic features, such as hedges, reporting verbs, directives, tense, have also been examined under these terms for the roles they play in stance construction (Crompton, 1997; Fløttum et al., 2006; Hyland, 2004, 2005; Salager-Meyer, 1994; Swales, 2004; Thompson, 2001).

Nevertheless, one linguistic feature salient in stance-making, namely the *Noun Complement* construction, where a stance head noun takes a nominal complement clause (as *fact that* in *The fact that catalysts exist in QRTs is not obvious*), has been under-researched. One notable exception is a corpus-based study by Charles (2007), which focuses on the variations in the use of stance nouns and propositions in the complement clauses between the disciplines of politics and materials science. The study demonstrates that differences in the choices of stance nouns and proposition sources reflect different disciplinary values. Other related major studies are Jiang and Hyland (2015) and Jiang (2017), in which the frequencies, forms, and functions of the *Noun Complement*

DOI:10.6814/NCCU202100169

construction were explored based on corpus analysis of research papers across different disciplines. Grounded in the same function-based classification of stance nouns, both studies found that academic writers not only widely drew on stance nouns and the ownership of stance (or in Jiang's terms, the "voice") to convey authors' epistemological views and judgments on subject knowledge, but they also used significant variations of such nouns and ownership to construct knowledge across different disciplines. These previous studies have also examined the pre-modification of stance nouns (e.g., the attributive possessive *Scheve and Slaughter's (2001)* in Example (1)) to reveal the effects of averral (personally taking responsibility for a position) and attribution (attributing it to another).

In contrast, little scholarly attention has been devoted to how the post-modification of stance nouns — the intervening constituent between stance noun and its complement clause, such as the expression *made by Nolte et al. (2017)* in Example (2) — is used to aver or attribute the propositions in the complement clause. More importantly, post-modification, given its varying grammatical forms from phrases to more complicated structures (e.g., relative clause), tends to occupy a broader variety of positions in a clause than pre-modification (usually in the form of single words or short phrases) to incorporate more meaning elements. Thus, it is reasonable to assume that it serves more linguistic and extra-linguistic functions than pre-modification, which are, undoubtedly, far beyond that of averral and attribution. As demonstrated in Example (3), the adjective phrase *common to many cell marking techniques* as the post-modification of the head noun *disadvantage* is actually performing the hedging function of mitigating the criticism against the *procedure* and of course the threat to the face of those who established it. Nevertheless, the body of corpus-based research on academic writing clearly exhibits a very real dearth of empirical studies into the forms and functions of post-modification in *Noun Complement* construction.

(1) Moreover, *Scheve and Slaughter's (2001)* **belief** that citizens tend to weigh adverse … (Jiang & Hyland, 2015: p. 545, emphasis in the original)

DOI:10.6814/NCCU202100169

(2) For this article, however, we accepted the assumption **made by Nolte et al. (2017)** [that all land transformations can be equated …]

[Social sciences]

(3) This procedure, however, has the disadvantage **common to many cell marking techniques** [that the cells selected for labelling need to be …]

[Natural sciences]

Moreover, prior research has revealed fundamental differences between disciplines in terms of the source of knowledge and the mode of knowledge construction. For instance, Becher and Trowler (2001) made a comparison between the "impersonal, value-free" "hard" sciences and the "personal, value-laden" nature of "soft" fields (p. 36). The former shows respect for the cumulative nature of knowledge progression, with new findings generally accredited by adding to the developments of the existing state of knowledge. On the other hand, the latter holds that "new knowledge follows altogether more reiterative and recursive routes" (p. 31) as writers put forward and revisit previously explored features of others' views so as to take a stance in relation to them.

Thus, the distinction between soft and hard sciences leads to considerable variations in stance construction in order to reflect the views or judgements of a particular discourse community (Charles, 2003; Jiang & Hyland, 2015). A large body of research has suggested that academic writers construct stances to reflect their ideologies and epistemologies as well as knowledge-building practices of their disciplinary communities. In other words, authorial stance is, to a great extent, discipline specific (Biber et al., 1999; Charles, 2003; Dressen 2003; Hyland, 1999a; Jiang & Hyland, 2015; Tucker, 2003). More importantly, a number of studies have proposed that stance construction can be complex, involving not only the choice of controlling verb, noun, or adjective but also the choice of proposition sources (as in the case of averral or attribution) (e.g., Charles, 2006; Hunston 1993, 1994; Thompson & Ye, 1991). Consequently, how writers construct a stance by using a post-modification structure may relate closely to the practice of a particular discourse community. Following this, there is no doubt that the various lexico-grammatical features

DOI:10.6814/NCCU202100169

of post-modification use call for a detailed comparison of its forms and functions across disciplinary fields.

Additionally, discourse, according to Kress (1989), is socially produced in particular communities and relies on them to make sense. Following this, texts in academic fields, as Bazerman (1993) notes, record writers' social practices and beliefs within their own disciplines, and embody the social negotiations of their disciplinary inquiry. In this sense, the writing of academic discourse is also subject to socio-culturally variable meanings and interpretations. As such, the lexico-grammatical patterns of stance-making in published research articles are not only discipline-specific but also influenced by a writer's first language and cultural background. Granted that "basically the L2 writer is writing from his or her own familiar culture" (Hyland, 2004: p. 47), it is not surprising that non-native English authors tend to draw on the *Noun Complement* construction in ways distinct from the mainstream members of the international academic community. Overall, the normed frequencies of the target construction are far lower in academic texts written by authors of non-English native languages than their native English counterparts. Specifically, their use of such a construction differs in the types of stance nouns (e.g., Hyland & Tse, 2005b), the complement proposition contents and sources (e.g., Parkinson, 2013), and the pre-modifications of stance nouns (e.g., Jiang, 2015). Nevertheless, little research has been undertaken to reveal the differences in the forms and functions of post-modification in *Noun Complement* construction used by writers with different first languages and cultural backgrounds.

Taken together, the present study therefore aims to explore how post-modification in *Noun Complement* construction is used differently between disciplinary fields and writers of contrasting native languages and cultures. Specifically, cross-disciplinary differences will be explored through the comparison of English-medium journal articles in branches of the natural (chemistry, physics, biology) and social sciences (applied linguistics, law, economics). To reveal the cross-cultural differences, all these articles are selected from the most influential journals published in the UK or US and in Taiwan or mainland China.

DOI:10.6814/NCCU202100169

Then, the methods proposed by Wood (2001) and Pan, Reppen and Biber (2016) will be adopted to ascertain a writer's first language. Briefly, native English writers will be operationally defined as any author who has a first and last name considered native to English-speaking countries, and is affiliated with an institution in a country where English is spoken as the first language. Chinese EFL writers will be any author affiliated with an institution in Taiwan or mainland China, while also having a first and last name considered native to native countries of Chinese (see Chapter 3 for full description of the inclusion criteria for L1 English and Chinese writers). Special attention will be paid to the differences in disciplinary and cultural use of post-modification in terms of its frequencies, forms, functions, and motivations based on two self-constructed corpora of these articles: the corpus of international journal articles (CIJA) and the corpus of Taiwanese/Chinese local journal articles (CCLJA) (for details, see Chapter 3 for description of the methods).

1.2 Overview of the study

This dissertation deals with a corpus-based and mix-methods investigation of post-modification in the *Noun Complement* construction in written academic discourse. It uses a corpus of research articles in the fields of natural and social sciences from international (in the UK or US) and local (in Taiwan or mainland China) English-medium academic journals. Cross-disciplinary and cross-cultural differences in the frequencies, forms, functions, and motivations of post-modification use were set as the focus of this study. This was determined because after examining previous investigations of the *Noun Complement* construction in academic writing, the post-modification features in such constructions were found barely touched upon by the research community. Also informed by the implications of findings in prior research, differences in the target construction were to be explored first in terms of its lexico-grammatical features along two primary dimensions of variation: variation by discipline, and variation by culture. These features were then analyzed, refined, and categorized through corpus-linguistic methods.

Following all these considerations, this dissertation examines four dimensions of the target structure. Therefore, the frequencies of various lexico-grammatical features of post-

DOI:10.6814/NCCU202100169

modification use were summed and compared in the work of (a) international and local writers in the natural sciences, (b) international and local writers in the social sciences, (c) international writers in the natural and social sciences, and (d) local writers in the natural and social sciences, with an aim of uncovering the preferred patterns in each group of writers. Here, local writers refer to those who meet the above operational definition of Chinese EFL writers and publish their work in English-medium journals based in the Great China region, while international writers are the contributors to the English language journals published in the UK or US and fit the operational definition of native English writers.

The comparisons in the four dimensions were made on a feature-by-feature basis from a quantitative perspective. Possible explanations for cross-disciplinary and cross-cultural differences in the frequencies and forms of post-modification use were provided based on the functions of the constructions. This is because language, as Halliday (2013: p. 15) proposes, "had evolved in the process of carrying out certain critical functions as human beings interacted with" their environment. Language itself needs to be seen as a resource for meaning-making instead of as a set of rules (Byrnes, 2009; Halliday, 1985, 1998; Halliday & Martin, 1993/1996; Halliday & Matthiessen, 1999). In this sense, the lexico-grammatical resources of language can afford users the opportunities to choose the most meaningful ones according to the functions required for the types of activity constituted by language (e.g., academic writing) (Schleppegrell, 2004). Additionally, the quantitative analyses were complemented by qualitative analyses which involved iterative comparative readings of all the coded instances of the post-modification features across the subcorpora to identify salient patterns and prototypical instances illustrative of the patterns.

In short, these four dimensions are expected to provide an integrated analysis of post-modification use in the *Noun Complement* construction, including its frequencies, forms, functions, and motivations. Moreover, they will further our understanding of how the *Noun Complement* construction can be used in the knowledge production and linguistic or extra-linguistic interaction between varying disciplinary domains and writers with different first

DOI:10.6814/NCCU202100169

languages and cultural backgrounds. In addition, they are also expected to provide insight into the areas that academic writers of non-English native languages can focus on to improve their use of post-modification in research articles.

1.3 Research questions

The following questions will be addressed in the present dissertation study:

1. Are there differences in the use of post-modification in *Noun Complement* construction between international and local academic writers in the natural sciences?
2. Are there differences in the use of post-modification in *Noun Complement* construction between international and local academic writers in the social sciences?
3. How, if at all, do international academic writers in the natural and social sciences differ in their use of post-modification in *Noun Complement* construction?
4. How, if at all, do local academic writers in the natural and social sciences differ in their use of post-modification in *Noun Complement* construction?

1.4 Outline of the study

Dissertation chapters are organized as follows. Chapter 2 presents an overview of literature from relevant research, with the focus on the studies that characterize variation in the use of the *Noun Complement* construction in varying disciplines and by writers of different native languages and cultures. It also provides the detailed accounts of the theoretical framework and analytical methods for studying post-modification use in the *Noun Complement* construction. Chapter 3 gives a detailed description of the methodology adopted to build and analyze the new corpus used throughout this dissertation. Chapter 4 demonstrates the cross-disciplinary and cross-cultural differences in the frequencies, forms and functions of post-modification in the *Noun Complement* construction. Chapter 5 discusses the reasons that underlie these variations, followed by the comparison between this study and prior research in terms of the lexico-grammatical features, functions and reason behind variations identified. Chapter 6 provides a synthesis of findings from the study chapters and makes suggestions for future research into the *Noun Complement* construction in academic writing.

DOI:10.6814/NCCU202100169

CHAPTER 2

Literature review

2.1 Introduction

This chapter first provides an overview of published studies on the *Noun Complement* structure as a nominal stance construction and identifies post-modification of the *Noun Complement* construction as an area little studied in academic writing; next, the chapter establishes the theoretical framework and analytical methods for studying this linguistic structure. Specifically, Section 2.2 reviews studies on stance nouns in academic writing and then into the *Noun Complement* structure as a nominal stance construction. The next two sections focus on the research into the disciplinary (Section 2.3) and cultural differences (Section 2.4) in the use of *Noun Complement* construction, attending respectively to its three main components of *stance noun, proposition in the complement clause* and *pre-modification*. These three sections will help establish the need for studying *post-modification*, the fourth component. Section 2.5 provides detailed accounts of the theoretical framework and analytical methods for studying post-modification structures in the *Noun Complement* construction.

2.2 Stance noun and its complement clause in academic writing

The use of stance nouns is quite common in academic written discourse (Charles, 2007; Coxhead, 2000; Jiang & Hyland, 2015), and thus they have attracted considerable attention in the literature. In early literature, Halliday and Hasan (1976) first identified a class of "general nouns" (p. 274), and demonstrated that they allow the writer to introduce "an interpersonal element into the meaning" (p. 276). Such nouns were then analyzed under a range of different names, including "unspecific nouns" (Winter, 1982), "anaphoric nouns" (Francis, 1986), "carrier nouns" (Ivanič, 1991) and "labels" (Francis, 1994). After examining a large amount of corpus data, Hunston and Francis (1999) proposed a possible new word class, namely, "shell nouns", whose use entails some kind of expansion in meaning in their immediate context (i.e., "lexical realisation", in Winter's terms (1977: p.7)) (as *premonition* in *He had an unshakable premonition that he would die,* where the *that-*

DOI:10.6814/NCCU202100169

clause expands the meaning of the noun *premonition*). They further categorized such nouns by something in the surrounding text that they refer to: namely, nouns that make reference to what is written or spoken (e.g., *announcement*), or to what is thought or believed (e.g., *doctrine*). Similarly, Schmid (2000) used a corpus of 225 million words from the Bank of English to identify and describe shell nouns from both a theoretical and a functional perspective. Shell nouns were distinguished according to three criteria: (a) they perform the semantic function by "characterising" chunks of information in clauses or longer texts, (b) cognitively, they allow writers to adopt the "temporary concept formation", which means writers can encapsulate information chunks in temporary nominal concepts (e.g., *awareness*, as in Example (4) below), and (c) concerning text organization, they link clauses or longer texts which "contain the actual details of information", thus serving a discourse-organizing function (p. 14). In addition, Flowerdew (2003) explored how "signalling nouns" have important discourse functions in establishing links across and within clauses, through making a comparison of their uses in lectures, journal articles and textbook chapters, and of their distributions across five disciplines.

(4) Of course, public recognition of the need to introduce a smog alert system may increase **awareness** that air pollution problems are very serious and that stricter long-term pollution control policies are needed to prevent smogs occurring in the first instance.

[NSS INT RA3]

However, many of these names proposed thus far suggest that researchers have mainly concerned about the discourse-organizing functions of these nouns, focusing on the way they serve as cohesive devices by "enclosing or anticipating the meaning of the preceding or succeeding discourse" (Aktas & Cortes, 2008: p. 3). The choice of head nouns, however, does more than simply link and organize discourse; rather it contributes significantly to the rhetorical construction of a writer's stance (e.g., Charles, 2003; Francis, 1994). As early as 1994, Francis notes that certain stance nouns can display an attitude, such as *advantage* and *difficulty*, in a way that marks either a positive or negative attitude towards the

DOI:10.6814/NCCU202100169

information they refer to. In terms of the attitudinally neutral ones, such as *fact* and *observation*, the attitudes signaled by them, however, rely on the context where they occur. Notwithstanding this, in the following decade, use of nouns to construct stance in academic writing had attracted little attention until Charles (2003) began to examine shell nouns in the sentence initial *This* + noun pattern (e.g., *This shortcoming* leads to problems …). In her study, she classified shell nouns into *metalinguistic* nouns (i.e., nouns used to establish relationships inside the discourse and to instruct readers on how to comprehend the linguistic status of a proposition; e.g., *distinction*, *claim*) and *non-metalinguistic* nouns (e.g., *result*, *observation*). She then demonstrated how the selection of such nouns empowers writers by allowing them to incorporate their own stances into texts in two disciplines (i.e., *politics* and *materials science*). In this way, writers actually orient their readers to how the information in the proceeding propositions is supposed to be understood.

More recently, the focus in this line of research has shifted from the stance noun itself to the stance noun and its complementary construction combined, in particular to the *Noun Complement* construction. The *Noun Complement* construction is a grammatical structure in which a head noun takes a "noun complement" either in the form of "*that*-clause", "*to-infinitive* clause", "*of* + *ing*-clause" or "*wh*-interrogative clause" (Biber et al., 1999: p. 645). Examples are given below.

(5a) There is a **danger** that the choice of condition may prejudge the structure under investigation, but with care such measurements can reveal much about the structure of the eddies.

[NSS INT RA2]

(5b) Hence it is also our **responsibility** to raise our voices for the rights of voiceless, powerless disabled persons.

[SSS INT RA5]

(5c) Indeed, an extra **advantage** of having more than 1 instrument in the main study is the "automatic" receipt of a very large calibration study.

[NSS LCL RA6]

DOI:10.6814/NCCU202100169

(5d) There is without a doubt much **controversy** over whether teachers can identify reliably those children with reading problems.

[SSS LCL RA9]

The shift to the *Noun Complement* construction can be attributed to the emerging consensus among a number of scholars that complement clauses are a highly important way that writers can foreground their stance to accompanying propositions in these clauses (i.e., *that*-clause, *to-infinitive* clause, *of* + *ing*-clause or *wh*-interrogative clause) through the selection of an appropriate head noun (e.g., Biber et al., 1999; Hyland & Tse, 2005a, b). To illustrate, *danger* in Example (5a) refers to the proposition in its complement "the choice of condition may prejudge the structure under investigation, but with care such measurements can reveal much about the structure of the eddies". Its content is thus specified in the *that*-complement clause. The complement proposition is seen as providing the semantic equivalence of the stance noun, so that the claim or information in the clause is what is being assumed or seen as potentially dangerous. In this way, the choice of such a noun foregrounds the writer's assessment of what follows and shows the reader how its content should be understood (Biber et al., 1999; Hyland & Tse, 2005a, b). The *to-infinitive* clause, *of* + *ing*-clause or *wh*-interrogative clause shares the same stance-making function.

Still, there are few studies that tend not to acknowledge the stance-making roles of the *Noun Complement* construction. Biber (2006b), for example, presented some illustrations of what he called *epistemic* nouns (showing certainty and likelihood; e.g., *principle*), *attitude/perspective* nouns (conveying attitudes and evaluations; e.g., *hope*) and *communication* nouns (distancing writers from the truth of the proposition or implying their different degrees of commitment; e.g., *news*) in his TOEFL 2000 Spoken and Written Academic Language Corpus (T2K-SWAL Corpus). Concerning the cases of these nouns followed by complement clauses, he concluded that even though the use of such a construction was found in the instructional/academic written registers (i.e., textbooks and course packs), they relied the least on stance features (meaning minimal stance-making roles). On the contrary, Charles (2007) also recognized the stance functions of the *Noun*

DOI:10.6814/NCCU202100169

Complement construction, by noting that this structure allows writers to comment on the proposition that occurs in the complement clause through choosing the stance-making head nouns (as *danger* in Example (4) above) of that clause.

Until quite recently, a series of corpus-based studies into this structure by Hyland and his associate (Jiang, 2015, 2017; Jiang & Hyland, 2015) have put forward the clear, thoughtful, thorough explanations of how the *Noun Complement* structure serves as a powerful and effective nominal stance construction. Specifically, they emphasized that head noun is not simply used to organize discourse or convey the details of information inside an empty shell (noun). Instead, it engages crucially in the rhetorical formulation of the writer's argument. It serves as an influential persuasive device, given that the choice of nouns foregrounds the writer's evaluation of the reliability of what follows, and demonstrates how the material needs to be understood to the reader. In other words, *Noun Complement* construction, by providing a range of stance choices (choice of nouns to convey certain stances), allows writers to build a clear stance at the outset in a way that brings readers into alignment with that stance, and thus a presumption is built on the reader of how to comprehend and interpret the proposition to be unfolded in the following complement clause. In so doing, readers are brought into the writer's preset point of view, an act which, in turn, promotes the construction and development of the writer's argument.

To summarize, since research conducted in this vein has shifted its focus from stance noun alone to the combination of stance noun and its complement clauses, many studies have provided a comprehensive picture of the uses, forms, and functions of the *Noun Complement* construction. More importantly, the broader and more profound understanding of the use of this structure in academic writing is usually developed through the cross-disciplinary and cross-cultural comparison in this line of research. Hence, in the next section, the studies into the use of *Noun Complement* construction across different disciplinary contexts will be reviewed first.

DOI:10.6814/NCCU202100169

2.3 Cross-disciplinary differences in the use of *Noun Complement* construction

A range of studies has explored the disciplinary differences in establishing stances through the *Noun Complement* construction (e.g., Charles, 2007; Hyland & Tse, 2005a; Jiang & Hyland, 2015; Jiang, 2017; Kim & Crosthwaite, 2019). A scholarly consensus has been achieved that the use of *Noun Complement* construction in stance-making practices are more frequent in the social sciences, given that the discursive and interpretative nature of the soft knowledge domains leads writers in these fields to more likely make a stance on the things that they talk about and evaluate both their own and others' pieces of work. Research conducted in this vein often focuses on three major components of the *Noun Complement* construction – stance noun, proposition in the complement clause, and pre-modification.

2.3.1 Stance noun

In the case of stance noun, studies tend to explore the disciplinary variations in the frequencies, forms and functions of these nouns, and attribute such differences mainly to the varying disciplinary community conventions that pertain to value systems and ideologies of the specific academic fields. In their research into evaluative *that* constructions [1] in the abstracts of research articles, Hyland and Tse (2005a) proposed that hard field writers more favored the nouns in the nominalized forms of research and reporting practices (as *finding* in *Cirrus HD-OCT high definition imaging provides agreement with the finding that the choroidal thickness is* …). This can be attributed to natural scientists' embrace of a core scientific ideology which highlights the importance of experimental work and the explanatory value of laboratory results. Moreover, the nouns with higher degrees of certainty (e.g., *demonstration*, *proof*) were drawn on far more in hard than soft disciplines, since the interpretive nature of the soft fields, where the challenging of others' interpretations occurs more often, entails the use of more tentative

[1] Evaluative *that* construction is a structure that allows authors to express evaluation and stance via the *that*-complement clause(s) in a super-ordinate clause (with noun, adjectival and verbal predicates), whose nominal forms (noun predicates) are, in turn, one kind of the *Noun Complement* construction (e.g., *We provide evidence that Shkl, pombe homolog of the STE20 protein kinase, can directly antagonize the Byr2 intramolecular interaction, possibly by phosphorylating Byr2.*).

DOI:10.6814/NCCU202100169

nouns (e.g., *suggestion*, *assumption*). This is supported in the follow-up research into evaluative *that* clauses by Kim and Crosthwaite (2019), where nouns that express doubt (e.g., *potential*, *suggestion*) were preferred by published academic writers in business studies, given their attempt to persuade the readers of their claims, which, in turn, also reflects the discursive and interpretative nature of the social sciences, where knowledge is built on suggestions, arguments, or assumptions. Similarly, Charles (2007) made a comparison between two contrasting disciplines (i.e., politics/international relations and materials science). Different from Hyland and Tse's (2005a) work, her study developed a systematic categorization of stance nouns, by dividing these nouns into five main groups (i.e., *idea*, *argument*, *evidence*, *possibility* and *others*). She then suggested that academic writers in politics/international relations drew on more nouns from *idea* (e.g., *idea*, *assumption)* and *argument* groups (e.g., *argument*, *contention*), inasmuch as their discipline constructs knowledge through examining ideas and building arguments, while also working at the creation of understanding. By contrast, their counterparts in materials science preferred the *evidence* group of nouns (e.g., *evidence*, *confirmation*), since knowledge in this field is advanced based on the evidence found in the experiments to accept or reject the hypotheses put forward.

In the follow-up studies by Jiang and Hyland (2015) and Jiang (2017), academic texts from more disciplines (e.g., applied linguistics, marketing, sociology, philosophy, electronic engineering, medicine, cell biology, physics) were incorporated into the corpora and stance nouns were categorized based on their functions as *entity* (orienting to objects, events, discourses or aspects of cognition; e.g., *paper*), *attribute* (concerning judgements and evaluations of the quality, status and formation of entities; e.g., *advantage*) and *relation* (elaborating how a writer understands the connection or relationship to information in a proposition; e.g., *difference*). In this regard, the *entity* category is further subdivided into *object* (e.g., *report*), *event* (e.g., *process*), *discourse* (e.g., *claim*), and *cognition* (e.g., *belief*) groups of nouns, and the *attribute* category into *quality* (e.g., *value*), *manner* (e.g., *method*), and *status* (e.g., *trend*) groups. The results showed that the soft fields used more *cognition*

DOI:10.6814/NCCU202100169

types and the nouns in the *attribute* category, whereas the hard sciences tended to draw on *event* types most frequently, with the non-use of *discourse* stance nouns. These patterns of use were supported by their findings in the most frequently used head nouns across disciplines. Specifically, *way* and *ability*, which belong to the *cognition* types and the *attribute* category respectively, were the most frequent nouns used by authors in soft fields, while their hard domain counterparts drew on *fact* and *evidence* most often, both of which are included in *event* types. The reasons they provided to explain these disciplinary differences in the use of *Noun Complement* construction, though in more conceptual and comprehensive manners, are rather similar to those adopted in Hyland and Tse (2005a) and Charles (2007) in nature, namely, different research practices and mode of knowledge construction across academic disciplines. Specifically, the three main reasons can be summarized as below:

Firstly, soft and hard fields have distinct natures, with soft areas being discursive, where "unclear intellectual boundaries normally make the knowledge webs so loosely-knit that identifying a credible problem is the main way for writers to justify their work to readers" (Jiang, 2017: p. 98), and hard disciplines being cumulative, where new claims are integrated into current knowledge, drawing on it as supporting testimony.

Secondly, soft and hard fields differ dramatically in the modes of knowledge construction, namely that in soft knowledge domains, knowledge is built on cognitive understanding and theoretical constructs, or in other words, soft knowledge fields tend to build knowledge on personal interpretation and intellectual negotiation, which are open to writers' evaluation and judgement. On the contrary, empirical evidence is the primary mode of knowledge construction in hard fields, or to be exact, knowledge in hard sciences is built far more on empirical evidence and the creation of facts through experimentation and observations.

Finally, soft and hard sciences are differentiated by their respective disciplinary value systems and ideologies. In the soft fields, "the fabric of established understandings has a wider weave" (Jiang, 2017: p. 99), inasmuch as knowledge problems are scattered, non-

DOI:10.6814/NCCU202100169

linear and reiterative, ranging over a broader stretch of academic and historical realms, and thus the development of supporting claims and warrants entails the innovation and plausibility of personal interpretation. Arguments are supposed to be explicitly interpretive and personal, which, in turn, needs to be carefully structured to convey explicit reasons and explanations. As a result, writers in soft disciplines are more liable to step in texts and make interpretative comments. On the contrary, hard scientists tend to avoid their intervention in texts and personal evaluation of both their own and others' work as a way to construct the impersonality and objectivity in their work. This is because knowledge progression in hard fields relies on new findings generally accredited by adding to the developments of the existing state of knowledge.

To conclude, although this line of research into stance nouns in the *Noun Complement* construction presents intriguing accounts of the frequencies, forms and functions of these nouns used across academic disciplines, the underlying reasons behind such variations are provided mainly from the perspective of disciplinary ideologies and conventions. As any specific discipline is situated in a wider social world, so is the writing in this field (Hyland, 2004), it is reasonable to propose that the discussion of the reasons that underlie disciplinary differences are far beyond the scope of the disciplines per se.

2.3.2 Proposition in the complement clause

This line of research also attends to the disciplinary differences in the propositions in the complement clauses. By investigating the source of the complement proposition, Hyland and Tse (2005a) found that hard field writers tended to avoid attributing the proposition in the *that*-complement clause to a human source (including the author or other researchers; e.g., as *They* in *They propose the theory that organizations become more homogeneous* …) more than their soft field counterparts. This is because hard knowledge is built on "non-contingent pillars of replication, falsification, and rigorous application of approved methods" (p. 55). On the other hand, writers in soft domains favored the abstract entity source (e.g., data, results) as a way to establish the legitimacy of their claims by shifting attention from those who undertake the evaluating or reporting process (e.g., the

DOI:10.6814/NCCU202100169

writers themselves) to the study itself (as *These results* in *These results support the view that social support may be associated with longer survival* …). Disciplinary differences also lie in what the proposition following the complementizer *that* referred to. Namely, the soft field exhibited its preferences for the complement propositions referred to the content of previous studies or research goals (as in *This study is designed to achieve the goal that the proposed methodology in decision analysis is executed* ...). Nevertheless, these differences were left unexplained, perhaps due to the limited explaining power of the factors that pertain only to modes of knowledge construction and disciplinary conventions in their study. Similarly, Charles's (2007) research also examined the proposition sources in the complement clauses. She proposed that in the field of politics/international relations, writers made much more extensive use of research sources (propositions advanced by other researchers or their work) and non-research sources (propositions expressed by non-disciplinary actors or entities) than their counterparts in materials science. This variation was also attributed to the disciplinary differences in research practices (e.g., political scientists need to refer to political figures and entities) and the construction of knowledge (e.g., emphasis on creation of understanding or interpretation in politics).

Different from the work of Hyland and Tse (2005a), Charles took a major step forward by investigating the combination of noun groups and the sources of their complement propositions (e.g., the *argument* group with research sources as in *Brenner and Spaeth's claim*), with the focus mainly on self- (the writer as the source of the proposition) and research sources. In the field of politics, the most frequent combination was the *argument* group with research sources, followed by the *idea* nouns with research sources, while materials science writers favored the *evidence* and *idea* groups with self-sources most. However, the underlying reasons given remained mainly the different constructions of knowledge between disciplines. Specifically, the mode of knowledge construction in hard fields entails building one's work on that of another. The combination of the *evidence* and *idea* nouns with self-sources thus provide writers the chance to show their own evidence and ideas in the complement propositions to be sound and reliable enough to be the basis

DOI:10.6814/NCCU202100169

for future cumulative building of knowledge in the fields. In the social sciences, given the loosely-knit knowledge webs, knowledge is normally constructed by presenting the views of other researchers in order to take up a stance in relation to them. Various stances may therefore co-exist simultaneously within the discipline. As such, using the combination of the *argument* and *idea* groups with research sources therefore leaves more room for writers in soft fields to position themselves with respect to other researchers and their work of the discipline, whereby the writers' stances can then be made towards the ideas or arguments of others in their specialist fields.

More recently, Kim and Crosthwaite (2019) drew on Hyland and Tse's (2005a) model to examine the disciplinary differences in the use of evaluative *that*-clauses, with the focus on the propositions in the complement clauses as well. Firstly, in terms of what the proposition following the complementizer *that* referred to, the most statistically significant differences between writers in business and medicine lie in the complement propositions referred to the content of previous studies (in line with Hyland and Tse's findings), followed by research goals. The underlying reasons behind these variations were still attributed to contrasting disciplinary modes of knowledge construction and community conventions between the fields of natural and social sciences. Specifically, the discursive and interpretative nature of the soft knowledge domains results in more evaluation of others' work, while medical writers prefer to implicitly evaluate others' findings in non-integral or non-reporting sentence structures rather than in evaluative *that*-clauses. Besides, the rigid 4-part BMRC format (Background-Methods-Results-Conclusions) of medical research article hinders writers from expressing research goals. These two groups of writers also differ widely on the source of the complement proposition. Medical writers usually chose to attribute the source of the complement proposition to an abstract entity, whereas their business counterpart did so to a human source (either the author or other researchers). This is mainly due to business writers' propensity to regard their or others' findings as the results of human interpretation in their specific field. By contrast, writers in medicine apply

DOI:10.6814/NCCU202100169

professional medical technologies to generating their results, whose presentation is thus based on the abstract sources in medicine (e.g., laboratory model).

To summarize, the reasons behind the disciplinary differences in the use of complement propositions in above studies are still offered based on a number of discipline-specific factors, not considered within a wider social context where the actual research work occurs.

2.3.3 Pre-modification

As to the component of pre-modification, Charles (2007) pointed out that nouns in the *evidence* group with self-source are much more pre-modified in materials science than the ones in either the *argument* or the *idea* group with research source in the field of politics. She further noted that *adjectives* and *quantifiers* as pre-modifiers can provide writers an additional opportunity to make their stances. For example, authors can express their approval for a stance noun by placing an adjective *satisfactory* in its attributive position. She also demonstrated disciplinary differences in taking a stance in relation to reader by soft field writers and to display objectivity by their hard field equivalence. That is, the former is to use the definite article to suggest that the information is already known to the readers (as *The* in *The notion that interstate competition* …), while the latter is to prevent the person marker from co-occurring with the stance noun (as *my* in *my contention that pressure from parts* …). She did not recognize these two lexico-grammatical features as pre-modification, to which they actually belong, given their placement in the attributive position of the stance noun. However, no reasons were stated to justify all these variations.

Jiang and Hyland (2015), compared to Charles (2007), treated the pre-modifying structures in a far more systematic way. They categorized them into three types of *first-person possessives* (e.g., *our* assumption), *attributive possessives* (e.g., *media's* claim), and *scholarly citation* (e.g., *Wikstrim's (2006)* assumption) based on whether this stance is averred as writers' own (averrals) or attributed to others (attributions). By applying this model, they found that hard field writers made more frequent use of first-person plural possessives, given the prevalence of multiple authorship in natural science articles. On the other hand, due to the particular conditions of production and interpretation of texts in soft

DOI:10.6814/NCCU202100169

domains (as noted above), their social counterparts more favored *scholarly citation*. Similarly, Jiang (2017) coded the categories as *overtly averred* (first-person possessives), *other human* (third-person possessives), *concealed* (avoid of any manifest possessive forms) or *abstract entity* (possessives of an entity or institution). He then found social scientists' more attribution of a stance to other humans, which was, in turn, attributed to their need to establish a "discursive and contextual framework" (p. 101) in the soft knowledge domain (see Section 2.4.2 for more details about the use of pre-modification).

Taken together, research conducted in this vein has provided a comprehensive picture of disciplinary differences in the frequencies, forms and functions of the three major components of the *Noun Complement* construction, namely, stance noun, proposition in the complement clause, and pre-modification. However, the cross-disciplinary analysis of another component of the target structure is missed in this picture – that is, post-modification. Any investigation of the *Noun Complement* construction should include or even be carried out in sequence of pre-modification, stance noun, post-modification and proposition in the complement clause. Since there has been no research investigating the disciplinary differences in the use of post-modification so far, the present study thus aims to explore how academic writers from different disciplinary fields use the post-modification component of the *Noun Complement* construction. Moreover, given the rich and varied grammatical forms of the post-modification structure from phrases to more complicated clausal constructions, this structure has potential to serve a wide variety of linguistic and extra-linguistic functions. This dissertation will develop an explanatory model based on the linguistic, rhetorical, relational, and socio-cultural aspects of how post-modification in the *Noun Complement* construction functions, with the aim of offering full accounts of the reasons that underlie the cross-disciplinary differences (for details of the explanatory model, see Section 2.5.5).

2.4 Cross-cultural differences in the use of *Noun Complement* construction

Studies into the use of *Noun Complement* construction by writers with different first languages (L1) and cultural backgrounds have also focused on the three main components

DOI:10.6814/NCCU202100169

of the target structure – stance noun, proposition in the complement clause and pre-modification.

2.4.1 Stance noun

The cross-cultural differences in the use of stance nouns in the *Noun Complement* construction were first reported in Hyland and Tse's (2005b) research into the evaluative *that*-clauses in the abstracts of research articles and Chinese ESL students' masters and doctoral dissertations. They noted that published writers were more liable to use stance nouns with expressions of doubt (e.g., *proposal*, *indication*) to convey meaning in a more tentative way. This was attributed to their greater familiarity with the more explicitly interpretative nature of the soft fields than student writers, thus attending more to the force of their claims and framing them as "suggestions, arguments, or assumptions" (p. 135). On the other hand, Chinese ESL students' inclination to choose more stance nouns that express affective and greater certainty is sometimes likely to cause their writing to look "overstated and perhaps rather anxiously persuasive" (p. 137). As such, the explanations provided for the cross-cultural differences in their study are still based on the narrow perspective of disciplinary ideologies and conventions. Parkinson (2013) adopted a similar corpus approach to compare the use of evaluative *that*-clauses in South African ESL undergraduate student reports and published research articles. In his study, L2 students were found to use a narrower range of stance nouns partly drawn from the informal and colloquial expressions characteristic of daily conversation (as *thing* in *I am not certain about the thing that if they knew more about it, the results would* …). This reflects the importance of the values in the social world outside university to them, whereby they attempted to draw on the kind of evidence and way of knowing characteristic of the social world but not academic writing. This, undoubtedly, offers support for analyzing the reasons behind the cross-cultural differences within a wider social context.

In another corpus-based comparative study of American and Chinese university students' argumentative essays, Jiang (2015), based on Jiang and Hyland's (2015) function-based classification of stance nouns, provided a more comprehensive picture of

DOI:10.6814/NCCU202100169

the differences between English native and EFL students' writing in terms of stance nouns. Specifically, Chinese EFL writers made much less frequent use of the *event*, *discourse* and *cognition* nouns in the category of *entity* (e.g., *evidence*, *argument*, *belief*) than their native counterparts. This implies that compared to L1 American students, EFL learners were reduced by their lower level of proficiency in English (e.g., limited academic lexicon) to a weaker position to define argumentation (with *discourse* nouns) and the world (with *event* or *cognition* nouns). Similarly, their reluctance to draw on the stance nouns referring to relations (e.g., *reason, grounds*) may offer them less chances "to establish discursive space for further elaboration and discussion" (p. 96). On the other hand, Chinese EFL writers tended to make most use of the nouns that indicate the stances towards the attributes of entities (e.g., *advantage*, *ability*), which, in turn, reflects their propensity to make their attitudinal evaluation and value-laden judgment. Additionally, EFL students showed a high reliance on a single type of stance nouns, by reusing a number of highly familiar vocabulary, which, thus, supports Parkinson's findings. Finally, it was comparatively infrequent for Chinese EFL writers to integrate the stance nouns from different categories or subgroups (i.e., the *event* and *discourse* or the *event* and *cognition* groups), in the way that they are placed in different sentences adjacent to one another respectively. In fact, such integration can aid writers in constructing strong arguments, for discoursal statements (from the *discourse* group) or cognitive points of view (from the *cognition* group) can be supported by factual events (from the *event* group) in nearby sentences.

Jiang also offered detailed explanations as to why Chinese EFL writers demonstrated above weaknesses, with emphasis on English competence and L1 influence. Specifically, Chinese EFL writers' limited academic lexicon and lack of awareness about academic register and genre conventions may contribute to their insufficient use of stance nouns in different categories. Moreover, "worldly unity" philosophy has a profound effect on the Chinese language, thus shaping its preferences for "generalized and holistic" words (Hu, 2005, p. 54), and prompting Chinese EFL writers to choose pragmatically vague vocabulary. Also, the *Noun Complement* construction is a typical English syntactic

DOI:10.6814/NCCU202100169

structure in which a head noun subsequently takes its specific lexical realisation, but in Chinese, lexicalization is placed before its head noun. Such a syntactic contrast may render Chinese EFL writers unfamiliar with this structure and reluctant to use it. Therefore, Jiang developed a more comprehensive understanding of the various influences on the use of the target structure between writers with contrasting first languages and cultural backgrounds than previous studies. However, his focus is too narrowly on accusing Chinese EFL writers of their low levels of proficiency in academic English, and the Chinese language of its idiosyncratic linguistic features, with much less emphasis on the rhetorical, relational, and socio-cultural aspects of how post-modification in the *Noun Complement* construction functions across cultures.

2.4.2 Proposition in the complement clause

Research into the *Noun Complement* structure as a nominal stance construction also examines the proposition in the complement clause, with the focus on the complement clause content and the source of the complement proposition. In Hyland and Tse (2005b), L2 student writers were found to be more enthusiastic about choosing their own findings, in particular the implications of the findings, as the contents of complement clauses than their expert counterparts. The reason for this is that students, as novice research writers, are in more pressing need of demonstrating the novelty, usefulness and disciplinary relevance of their work in order to convince other members of their academic fields to take their findings seriously. On the other hand, however, they, especially those as soft field writers, were more reluctant to give more space to the work of others in the *that* clause. This can be attributed to the reason that they are not attending to the discursive and overtly negotiable aspects of argument in the social sciences. Differently, L2 student writers in Parkinson's (2013) study tended to adopt the common knowledge that people in general hold or the religious beliefs (as in *Legislators would not legalize euthanasia since they hold the belief that it is against God's wishes* …) as the content of the *that* clause. This indicates that the student writing is not grounded in the academic world, and that it is, similar to their use of the stance nouns inappropriately informal and colloquial for academic writing,

DOI:10.6814/NCCU202100169

another conversational means of offering evidence. Moreover, they were disposed to incorporate others' work into the proposition in the *that* clause without giving credit to the original authors, or in other words, with the origin of the information unclear from context. This demonstrates their unawareness of academic norms, or to be precise, that they were in the early stages of acquiring the academic value of scrupulously attributing sources to the quotations. In addition, novice undergraduate writers also favored participants' opinions as the *that* clause contents, whereas their expert counterparts inclined to choose the model construction and their contributions to the field as the contents. This difference thus reflects L2 writers' unfamiliarity with the purposes they need to envisage for research work.

As to the source of the complement proposition, L2 students in Hyland and Tse's investigation displayed their stronger preference for the attribution of them to abstract entity sources. This is because many students view the use of non-human sources as a way of showing their academic competence. In so doing, they are able to emphasize their unique role (not with the help of other researchers) in putting forward the proposition in the *that* clause. In the similar vein, Parkinson (2013) found that it is most common for L2 student authors to attribute the complement propositions to the abstract entity sources, which was similar to the pattern of use in published research articles. This is likely to prove that L2 undergraduate students had in effect been moving towards the acquisition of academic language and values. Notwithstanding this, when they chose themselves as the sources of the *that* clause contents (i.e., attributing the source to the author), they tended to use pronouns (i.e., *we*, *you*, and *one*) in making their claims (as *You* in *You can find the fact that some people even never heard about it* …). This was done because they intended to emphasize their feelings and beliefs. All these, however, indicate that such student writers were not yet accustomed to the discourse norms of their disciplines or of academic writing in general. Again, the reasons given above are all centered around disciplinary ideologies and conventions.

DOI:10.6814/NCCU202100169

2.4.3 Pre-modification

In terms of premodification use, only Jiang's (2015) research touched upon non-native student writers' overuse of attitudinal evaluation and self-mention through the adding of pre modifiers (as in *wonderful chance* and *our own responsibility*), which is likely to defy readers' expectation for English argumentative essays, and thus to cause their skepticism about the argument's credibility. The underlying reasons of this included these EFL writers' lack of competence in academic English, the L1 (Chinese) influence on them, and the misinformed writing instruction by English teachers in China. However, the adding of accusing English language teachers in China of their improper instruction (e.g., encouraging students to use "generalized and holistic" words) does not improve the situation that the analysis of reasons that underlie the cultural differences in the study are restricted to the two factors of English competence and L1 influence (see Section 2.4.2 for more details about the use of pre-modification).

Taken together, although research into cross-cultural differences in the use of *Noun Complement* construction has also built up a detailed picture of the frequencies, forms and functions of its three major components, the component of post-modification is still largely missing in this picture. Against this backdrop, the present study is thus intended to investigate the use of post-modification to stance nouns by academic writers with different first languages and cultural backgrounds. Moreover, an explanatory model built on the linguistic, rhetorical, relational, and socio-cultural aspects of how post-modification in the *Noun Complement* construction functions will also be drawn on to account for cross-cultural differences beyond the factors of English competence and L1 influence. The next section will introduce three bodies of theories this dissertation has drawn on to develop its scope of study and analytical framework.

2.5 Theoretical framework and analytical methods for studying post-modification structures in *Noun Complement* construction

This dissertation draws on three bodies of theories to develop its scope of study and its analytical framework. Biber and Clark's (2002) and Biber and Gray's (2016) typological

DOI:10.6814/NCCU202100169

theories of present-day patterns of nominal post-modification in English will provide a systematic means of identifying the lexico-grammatical features, functions and motivations of general noun post-modification use. However, given the use of the target structure beyond the scope of Biber and his associates' theories, the analytical approach taken in the studies into pre-modification in the *Noun Complement* construction will be included to equip this study with a narrowed down framework. Lastly, as the functions of general noun post-modifiers in Biber and his associates' theories are mostly linguistic (syntactic-semantic) and pre-modification in the *Noun Complement* construction mainly performs functions pertaining to the voice of ownership of a stance or writers' rhetorical effort, this dissertation also draws on theories about language functions to understand linguistic and extra-linguistic functions performed by the target structure.

2.5.1 Typological theories of present-day patterns of nominal post-modification in English

Biber and Clark's (2002) and Biber and Gray's (2016) typological theories of present-day patterns of nominal post-modification in English are developed from the detailed account of noun modification in the *Longman Grammar of Spoken and Written English* (Biber et al., 1999). Table 2.1 illustrates the major postmodifier types of noun phrases.

Their theories establish two main patterns of nominal post-modifiers used in academic writing. First, phrasal grammatical features functioning as NP post-modifiers (i.e., prepositional phrase and *of*-phrase) have increased in use over time. The heavy reliance on

Code	Form	Example
1	Relative clause	the penny-pinching circumstances *that surrounded this international event*
2	*Ing*-clause	the imperious man *standing under the lamp-post*
3	*Ed*-clause	a stationary element *held in position by the outer casting*
4	Prepositional phrase	compensation *for emotional damage*
5	*Of*-phrase	this list *of requirements*

Table 2.1 Major postmodifier types extracted from the typology of present-day patterns of noun post-modification in English (Biber & Clark, 2002: p. 46)

DOI:10.6814/NCCU202100169

of-phrases even forces the researchers to distinguish between *of*-phrases and other prepositional phrases. Second, dependent clauses have declined in frequency when they function as noun post-modifiers. Based on the findings of these two patterns, they challenge two stereotypical theoretical assumptions in prior linguistic studies: (a) that grammatical complexity equals structural elaboration, realized in particular through the denser use of dependent clauses, and (b) that grammatical innovations do not happen in academic written discourse. Their theories demonstrate that embedded phrasal structures (e.g., noun post modifiers) are as important as embedded dependent clauses to the entire system of grammatical complexity, which, in turn, makes the use of post-modification in the *Noun Complement* construction, whose complexity features are both phrasal and clausal, worthy of investigation.

Moreover, post-modification in the *Noun Complement* construction is linguistically innovative in research articles as well, given that it is not so common a practice among writers to place such constructions between a head noun and its complement clause. This is because, seen from a cognitive processing perspective, post-modification construction is disruptive since it must be processed before reaching the complement clause of the head noun. Semantically, it breaks apart the proposition expressed by the joint effort of head noun and complement clause. As a result, this pattern of use is likely to overlook readers' processing need and hinder their understanding of the sentence meaning. However, although instinctively one would in the first instance expect post-modification in the *Noun Complement* construction to be barely used, it is in fact widely popular in the stance repertoire of academic writers in various disciplinary domains and with different first languages and cultural backgrounds. This is also supported by Biber and his associates' theories, which, similarly, suggest that even though a grammatical form (e.g., relative clause) may intuitively seem to be most used as nominal post-modification, the uses of other forms (e.g., prepositional phrase) are in fact equally or more common.

Their theories also account for the underlying motivations for these changes, among which the most important one is an informational purpose (i.e., conveyance of

DOI:10.6814/NCCU202100169

informational content). Actually, phrasal and clausal grammatical features in academic written discourse are significantly associated with grammatical complexity. Specifically, what makes academic writing grammatically complex is linked with structural elaboration and embedding of clauses and phrases from the traditional linguistic perspective. In this sense, the increased use of embedded phrasal structures (e.g., nominal post-modification) indicates a growing tendency towards phrasal complexity wherein clauses are maximally compressed into complex phrases. This is done for purposes of presenting more informational content on an efficient and concise basis, or in Biber and Gray's (2016) terms, of "convey[ing] the maximum amount of information in the fewest words possible" (p. 207). It is because academic sub-disciplines have proliferated and become increasingly specialized over the past century, which, in turn, leads to the "information explosion" (p. 50), and which, as a result, asks for the "economy of expression" (p. 129). Following this, the present study can also reveal whether phrasal structural devices functioning as NP post-modifiers in the *Noun Complement* construction outperform their clausal counterparts in frequencies, due to the similar functional and cultural reasons.

The second factor behind these changes is a specialist readership. Generally, phrasal complexity features are less explicit in meaning than their clausal counterparts, granted that they omit certain structural elements (e.g., predicates of different types; see Halliday, 1979; Halliday & Martin, 1993/1996), as a way to compress information into dense constructions. Since most of the readers of these academic research articles are also specialists in the same field, they usually share extensive expert background knowledge with the writers, which, therefore, prompts a substantial loss of explicitness in specifying the writer's intended meaning. As such, modern science writing has been witnessing the development and prevalence of phrasal complexity features since the turn of the twentieth century.

Production circumstances, as the last factor, also contribute to the ebb and flow of complexity features. Specifically, academic text is not produced in real time but, on the contrary, it allows careful and deliberate production through preplanning, editing and

DOI:10.6814/NCCU202100169

revision on the part of writers. As such, this makes it possible for them to take time in the construction of such maximally compressed structures (e.g., extensive nominal post-modification as in this case), including the involving extra-linguistic functions performed through them.

More importantly, their theories entertain the notion that academic research writing has underlying lexico-grammatical features that serve particular functions. Deeper consideration by them reveals that at a more fundamental level, occurrence of linguistic features is always meaningful and functional in academic written texts. In terms of nominal post-modifying features, they highlight a wide variety of differences in function and meaning through the use of nominal post-modifiers by academic writers, particularly the meaning relations with the head noun signaled by the post-modifying device (e.g., prepositional phrase). Specifically, the range of such meaning relations can be classified into two major types based on whether it is used to express concrete/locative meanings (*declining incidence rates in men*) or abstract meanings (*recent work on public sphere theory*). In fact, around 60% of these constructions in academic texts are used to convey abstract meanings rather than the other ones.

The post-modification, in the cases of concrete/locative meaning expression, normally serves to (a) express genitive meaning relationships (*the chairman of a major US corporation*), (b) convey locative meanings of different kinds (e.g., geographic location as in *most call centers in Australia* or textual location as in *the invented examples in this section*), and (c) supply additional descriptive information on the head noun, or form a co-referential relationship with it, in both of which the two constituents are usually separated by commas or parentheses (*two European countries (Finland and France)*; *the Pentagon memo, an internal document, was attached*). The abstract uses, in contrast, not only perform above functions but also provide strong evidence for the functional expansion of noun phrase post-modification. To be exact, such uses include (a) identifying topical domain (*problematic concepts in medicine*), (b) marking post-modifier as the semantic patient (the person or thing affected or acted upon by the action expressed by a verb, as *car*

DOI:10.6814/NCCU202100169

in *Paula fixed the car*) of the process described by the head noun (*a decrease in military spending*), (c) describing abstract state referred to by the head noun existing at the same time or place as the process in the post-modification (*difficulty in understanding what a discipline is*), (d) identifying the purpose of the entity referred to by the head noun (*a formula for testing the grade-level difficulty of reading materials*), and (e) conceptualizing head noun as the evidentiary basis of an argument, with the post-modification representing that argument (*grounds for concluding an innate sense of self*).

Taken together, the present study will draw on Biber and his associates' typology of present-day patterns of noun post-modification in English, as a means of identifying the lexico-grammatical features of the post-modifications used in the *Noun Complement* construction. In so doing, a comprehensive picture of the forms of the target structure will be pieced together. Also based on the concrete and powerful illustration of the functionality of noun phrase post-modifiers in their theories, how different lexico-grammatical features of the post-modification structures fulfil their functional roles in academic written texts will be systematically examined. Last but not least, an attempt to reveal the underlying reasons for the variations in the frequencies, forms and functions of the target structure will be made. This will be done with the help from the comprehension and interpretation of the textural, functional and pragmatic motivations shared by academic writers to drive towards various grammatical forms of noun post-modifying devices in their theories.

2.5.2 Use of pre-modification in *Noun Complement* construction

Given that the theories advanced by Biber and his associates do not touch upon post-modifying structures in the *Noun Complement* construction, this dissertation will also draw on the analytical and interpretative findings of pre-modification use in the *Noun Complement* construction in academic written English (see Table 2.2 below).

Firstly, the placement of pre-modification in the attributive position of the head noun significantly restricts its lexico-grammatical features into adjectives and possessives of

DOI:10.6814/NCCU202100169

Study	Disciplinary field	Writer background	Lexico-grammatical feature
Charles (2007)	a social science (politics/international relations), a natural science (materials science)	English native expert writers	adjective, quantifier
Jiang (2015)	college argumentative essays	English native university student writers, Chinese EFL university student writers	adjective, possessive
Jiang & Hyland (2015)	applied linguistics, marketing, sociology, philosophy, electronic engineering, medicine, cell biology, physics	published academic writers	possessive
Jiang (2017)	humanities, social sciences, political law, medicine, technical engineering, natural science	published academic writers	possessive

Table 2.2 Overview of previous corpus-based studies on pre-modification in *Noun Complement* construction in academic texts

different kinds. Charles (2007) identified *adjectives* and *quantifiers* as the pre-modifying grammatical forms most often occurring with stance nouns, but without subdividing them into a number of further groups. Jiang and Hyland (2015), by contrast, paid exclusive attention to possessives, by categorizing them into three major types – *first-person possessives* (e.g., *our* assumption), *attributive possessives* (e.g., *media's* claim), and *scholarly citation* (e.g., *Wikstrim's (2006)* assumption). Similarly, Jiang (2017) also focused on just the use of possessive pre-modification. Differently, he classified them as *overtly averred* (first-person possessives), *other human* (third-person possessives), or *abstract entity* (possessives of an entity or institution). In another research of Jiang (2015), two lexico-grammatical features of *adjectives* and *first-person possessives* were identified.

DOI:10.6814/NCCU202100169

Although the use of pre-modification in the *Noun Complement* construction displays quite limited lexico-grammatical features (adjectives and possessives), the ways in which the categorization schemes were created by above researchers still give useful hints on the typology of the lexico-grammatical features of post-modification in the *Noun Complement* construction. Specifically, the categorization in the present study will then be developed based on the feature categories of both general grammatical forms (e.g., prepositional phrase) and more specific subcategories (e.g., subdividing the lexico-grammatical feature of *punctuation* into *comma*, *colon*, *dash*, and *semi-colon*). Also, the feature categories of pre-modification in the *Noun Complement* construction can aid in identifying the corresponding features of its post-modifying counterpart (e.g., possessive nouns as pre-modifiers (*Erikson's* theory) are equivalent to *of*-phrases (the theory *of Erikson*), *ed*-clauses (the theory *proposed by Erikson*), or relative clauses (the theory *that Erikson proposes*) as post-modifiers).

Secondly, the functions of pre-modification in the *Noun Complement* construction are largely centered around the ownership of the stance noun, with certain studies touching on the extra-linguistic roles played by the target structure. Charles (2007) proposed that the use of pre-modification is thought to offer writers an additional chance to make their stances. For example, writers can show their approval to a stance noun by placing an adjective *satisfactory* in the pre-modification position of that noun. Provided such evaluative adjectives are used to undertake the positive assessment of writers' own pieces of work (e.g., the present study reports the *interesting* findings that …), they are more likely to go unchallenged by other members of their research communities. Moreover, she also demonstrated the function of the definite article as the pre-modifier to suggest that the information is already known to readers (as *The* in *The notion that interstate competition* …), which, in turn, reflects writers' construction of a stance with respect to the reader. On the other hand, writers can avoid using the person marker as the pre-modification to project an objective stance (as not using *my* in *my contention that pressure from parts* …).

DOI:10.6814/NCCU202100169

Based on their categorization of three major types of possessive pre-modification, Jiang and Hyland (2015) revealed the functions of such possessives in conveying writer attitudes. Specifically, first-person possessives are placed in pre-modification position, when writers personally take responsibility for the stance expressed through the head noun (as averrals), whereas if it is attributed to others (as attributions), either attributive possessives or scholarly citation are employed. Similarly, the functional roles that possessive pre-modification has been found to play in Jiang's (2017) investigation are indicating whether the stance expressed through the head noun is averred as writers' own or attributed to others. In fact, Jiang's functional categorizations of possessive pre-modification in the *Noun Complement* construction, to a large extent, collapse, separate and reorganize the categories in Jiang and Hyland's (2015) model. Specifically, Jiang drops *scholarly citation*, whose formats mainly include placing the names of cited authors in possessive pre-modification positions (e.g., *Sartre's* theory), and shifts its function to a new category of *other human*. *Attributive possessives* are too general to refer precisely to either human or non-human possessions and thus separated into *other human* and another new category of *abstract entity*. *First-person possessives* is recoded as *overtly averred*, with its function unchanged.

Jiang (2015) mainly highlighted the rhetorical aspects of how pre-modifying structures in the *Noun Complement* construction function. Adjectives are found to play the functional role of showing writers' attitudinal feelings (e.g., *wonderful* chance), whereas the function of first-person possessives is to establish an overt authorial presence in the text (e.g., *our own* responsibility). The above categorizations of pre-modifying functions provide useful references for designing the categorization scheme of post-modifying functions in the *Noun Complement* construction, in particular those concerning the voice of ownership of stance noun and the rhetorical effort made by academic writers.

Finally, research conducted in this vein also reveals the underlying motivations and enabling factors behind different uses of pre-modification in the *Noun Complement* construction across different disciplinary and cultural contexts. In Charles's (2007) study,

DOI:10.6814/NCCU202100169

the researcher indicated the disciplinary differences that nouns in the *evidence* group (e.g., *evidence*, *confirmation*) with self-source (writer as the source of the proposition in complement clauses) are much more pre-modified in materials science than the ones in either the *argument* (e.g., *argument*, *contention*) or the *idea* group (e.g., *idea*, *assumption*) with research source (other researchers or their work as the propositional source) in the field of politics. However, the underlying reasons behind these differences are not investigated, since pre-modification is not regarded as a powerful and effective linguistic and rhetorical device on its own in Charles's study, but rather an accompanying modifier subordinate to its head noun.

On the contrary, Jiang and Hyland (2015) acknowledged the key linguistic and extra-linguistic roles that pre-modification in the *Noun Complement* construction plays in stance-making practice across academic disciplines. They attributed hard field writers' denser use of first-person plural possessives to the prevalence of multiple authorship in natural science articles. On the other hand, their social scientific counterparts demonstrated a stronger preference for *scholarly citation*, given the particular conditions of production and interpretation of texts in soft domains. Namely, in light of the vast literature open to greater interpretation, the key findings more heavily borrowed from related fields, and the inconsistent and unclear criteria for supporting and denying claims, readers of social science journals are usually not assumed to have the same interpretative knowledge, whereby writers need to establish and elaborate a context through citations. Notwithstanding this, there was a general trend among writers in both natural and social scientific modes of inquiry towards advancing their arguments implicitly or attributing them to others. Taken together, this research succeeds in providing a more comprehensive picture of the reasons behind the use of the target pre-modifying structure.

Disciplinary differences, in Jiang's (2017) study, mostly lie in social scientists' more attribution of a stance to other humans through possessive pre-modification in the *Noun Complement* construction (e.g., *Hume's* demonstration that induction cannot be justified by …). This can be attributed to the need to establish a "discursive and contextual

DOI:10.6814/NCCU202100169

framework" (p. 101) in the soft fields. To be more precise, through the incorporation of prominent disciplinary scholars' names into their work, writers in social sciences are able to form an intellectual alignment with key members of their disciplinary community in a way that facilitates the process of discussing and interpreting their ideas and arguments from a new perspective. His study also concludes that published academics in both fields are more liable to draw on the possessive pre-modification in the form of *other human* and *abstract entity*, given their reluctance to "baldly present a personal stance" (p. 102). By adopting a clear functional approach and being consistent in his categorizations, Jiang's research provides a powerful analytical tool for revealing the underlying reasons attributing to the differences in the use of the target structure across specialist areas.

Different from the above research into the target pre-modifying structure, Jiang's (2015) work exhibits the differences in this pattern of use between writers with contrasting first languages and cultural backgrounds (American and Chinese university students). Specifically, Chinese EFL writers expressed too much personal affect through the deployment of adjective pre-modification in their work, while their overuse of first-person possessive pre-modification also results in unnecessary author visibility in the genre of academic writing. This is mainly due to these EFL writers' lack of competence in academic English, the L1 (Chinese) influence on them, and the misinformed writing instruction by English teachers in China. Specifically, their limited English academic vocabulary may bring the informal and colloquial expressions characteristic of daily conversation into their academic texts, thereby rendering the occurrence of emotional expressions much more possible. This could also relate to their lack of awareness about English academic register and genre conventions. Additionally, writing instruction in the English language classroom in China tends to lay indiscriminate emphasis upon personal expression of attitudinal evaluation. Non-existence of the *Noun Complement* construction in the Chinese language is likely to cause their inappropriate use of adjective and possessive pre-modification in the *Noun Complement* construction as well. As a result, such overuse is liable to contradict

DOI:10.6814/NCCU202100169

readers' expectation for academic written texts, which might, in turn, cause readers to question the credibility of writers' claims.

To summarize, the analytical approach taken in the above studies to investigate pre-modification in the *Noun Complement* construction, compared to Biber and his associates' theories, provides this dissertation a more specific and pertinent framework to look at the uses, forms, and functions of post-modification structures. Although the exploration of pre-modification in the *Noun Complement* construction barely gives useful hints on the typology of linguistic features of nominal post-modification, this construction performs more similar linguistic and extra-linguistic functions than other noun post-modification not within the *Noun Complement* construction. In fact, the only difference seems to be whether the modifier is placed in front of or after the head noun, thereby rendering the reference to the analytical approach in this line of research valuable. Additionally, if it is for identifying the forms of post-modification structures, this dissertation can always refer to Biber and his associates' typological models, although it is not designed especially for nominal post-modifiers in the *Noun Complement* construction.

2.5.3 Theories about functions of language

As noted above, general noun phrase post-modifiers mostly serve linguistic functions, or, to be more precise, syntactic-semantic functions: namely, communicating a number of different meaning relationships with their head nouns (e.g., expressing genitive, locative, or identity meaning relationships), or to convey varying semantic associations with them (e.g., identifying topical domains, describing abstract states). The functions of pre-modification in the *Noun Complement* construction, though more parallel to its post-modifying counterpart, are reduced to two main facets: the voice of ownership of stance noun and rhetorical effort made by writers. This is mainly due to the placement of such premodifiers in the attributive position of the head noun, which, in turn, offers writers a restricted range of lexico-grammatical options. Therefore, this dissertation also draws on theories about the functions of linguistic features, with the aim of providing more comprehensive accounts of

DOI:10.6814/NCCU202100169

the linguistic and extra-linguistic functions performed by post-modification in the *Noun Complement* construction.

Leech (1981, 1995), from semantic and discursive points of view, develops his classification of five functions of language. Specifically, the *informational* function is considered neutral and most important, given its role in purely conveying informational content, while the ways in which an originator's (e.g., an author's) feelings and attitudes are expressed demonstrate the *expressive* function (e.g., expression of affective meaning). A third function of language is the *directive* function whose aim is to influence the behavior or attitudes of receivers (e.g., readers). The fourth function of keeping communication channels open and social relationship in good repair is coded as the *phatic* function. This, along with *directive* and *expressive* functions, is considered to be most directly related to the social roles of language. Lastly, the *aesthetic* function refers to "the use of language for the sake of the linguistic artefact itself, and for no ulterior purpose" (1981, p. 41), with its emphasis on creating an artistic effect (e.g., poetic use of language). Leech acknowledges the inseparability of certain functions (e.g., *directive* and *expressive*), which, thus, allows the combined fulfilment of several different functions. Following this, a clause is likely to be read as at once informational, expressive, and directive. Lyons (1995) further simplifies Leech's five-part model by proposing a dichotomous classification of language functions. The *descriptive* (or *propositional*) function is defined as the use of language to convey the content (referential meaning) of statements that can be evaluated as true or false. On the other hand, the *non-descriptive* (or *socio-expressive*) function refers to the use of language to communicate expressive meaning that is tied to attitudes, emotions, or feelings of discourse participants.

From a systemic functionalist perspective, Halliday (1973, 1978, 1985) proposes that there are three main functional components or *metafunctions* reflected in the morphosyntactical patterning in language. The three components are: *ideational* (including *logical* and *experiential*), *interpersonal*, and *textual*. The *ideational* metafunction of language relates to the way that language functions as a model of reality, or, to put it

DOI:10.6814/NCCU202100169

another way, as a model of the material and social world where we live, and the world of our own consciousness (including perception, emotion and imagination). It can be further split into an *experiential* and a *logical* metafunction. The former, whose grammatical construing is accomplished principally through the coding of processes (actions, states, or relationships), deals with the representation and construction of what we talk about. The latter, otherwise, concerns the use of language to construe logic relations (dependency and interdependency relations) between the two things under discussion (e.g., the head noun and its embedded clause being in a dependency relation, given their unequal status). The way in which language serves to establish and maintain social relations is reflected in the *interpersonal* metafunction, which is, to be exact, concerned with expressing our own attitudes and judgments, while also seeking to bring others into alignment with them. This function is mainly realized in grammar through the resources of mood (e.g., statements, questions, demands, etc.) and modality (e.g., modal verbs and adverbs) in a language.

In fact, the first two metafunctions (i.e., *ideational* and *interpersonal*) orientate towards the first-order reality whose existence is independent of language. They are usually put into effect under the extra-linguistic circumstances, namely, the material and social world in which human relationships are primary. In contrast, the *textual* metafunction is oriented to the second-order, symbolic reality which is brought into existence by language per se. In other words, it is intrinsically linked to language itself, whereby it concerns the use of language to make links with itself and with the situations in which it is used. Its role is also an enabling one, in the way that it "serves to enable the presentation of *ideational* and *interpersonal* meaning as text in context" (Matthiessen & Halliday, 1993: p. 13). Figure 2.1 below diagrammatically demonstrates the relationship between language metafunctions and orders of reality.

It is also worth noting that these functions or metafunctions, from the systemic linguistic point of view, are not just perceived as characteristic of utterances or texts, but rather as an intrinsic and essential organizing principle of the language system, particularly of the lexico-grammatical patterning of linguistic structures and forms. In this sense,

DOI:10.6814/NCCU202100169

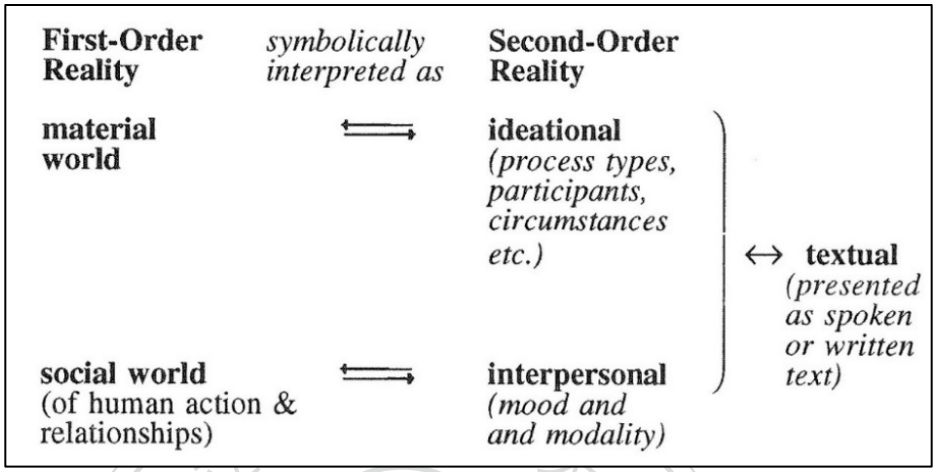


Figure 2.1 Metafunctions of language and orders of reality (from Shore, 1993, following Matthiessen & Halliday, 1993: p. 189)

granted that the notion of function is fundamental to the composition of the words and grammatical organization in a language, a clause, whose grammatical structure comprises a number of different sentential constituents, each of which may correspond to a specific function, can concurrently serve a number of different functions. This thus supports the multi-functionality theory in Leech's model from a novel interpretative angle. However, Halliday degrees with Leech and other linguists on viewing the function of conveying informational content (e.g., Lyons' *descriptive* function) as the most important function of language. Instead, it is only one of several functions that underlie the grammatical organization of a language.

Biber (1991) broadened and developed Halliday's tripartite model of metafunctions into the typology of seven major functions that can be fulfilled by lexico-grammatical features (see Table 2.3). The first two and the fourth functions in Biber's classification are parallel to Halliday's, even with the category names unchanged, but what needs to be noted here is that, in Biber's model, the ideational function refers specifically to the way that grammatical forms are used to express propositional or referential meaning. As such, both his *ideational* and *textual* functions are intrinsic to language itself, with the former dealing

DOI:10.6814/NCCU202100169

I. Ideational functions

A. Presentation of propositional meaning
B. Informational density

II. Textual functions

A. Different ways of marking informational structure and prominence
B. Different ways of marking cohesion
C. The extent to which informational structure, prominence, and cohesion are marked

III. Personal functions

A. To mark group membership of addressor
B. To mark idiosyncratic characteristics of addressor
C. To express attitudes towards the communicative event or content

IV. Interpersonal functions

A. To mark role relations
B. To express attitudes towards particular participants

V. Contextual functions

A. To mark physical or temporal setting
B. To mark purpose
C. To mark the psychological 'scene'

VI. Processing functions: caused by or in consideration of the production and comprehension demands of the communicative event

VII. Aesthetic functions: personal and cultural attitudes towards form

A. To conform to grammatical prescriptions
B. To conform to "good style"

Table 2.3 Functions of linguistic features (from Biber, 1991: p. 35)

with clausal structure and the latter text-internal structure, respectively. Notwithstanding this, Biber supports Halliday's opinion that the function in the conveyance of informational content should not be considered as the primary function of language. In this sense, this dissertation will also examine whether the function with an informational focus plays the most vital role in the use of post-modification in the *Noun Complement* construction.

The rest of the functions are oriented towards the extra-linguistic reality, wherein linguistic structures can be drawn on to mark information beyond the text itself. Specifically, the newly developed *personal* function is concerned with marking a speaker

DOI:10.6814/NCCU202100169

or writer's group membership, personal style and attitudes towards a communicative event or message content, whereas the relationship between discourse participants is most central to the contrasting *interpersonal* function. The aspects of such a relationship usually include role relationships (e.g., power and status between writer and reader), clearly conveyed attitudes towards participants (e.g., readers), the shared knowledge base between participants, and the likelihood of interaction in communicative events. In this sense, Biber's subdivision of the function that concerns discourse participants into *personal* and *interpersonal* functions refines Halliday's definition of the *interpersonal* metafunction, by attending to more detailed aspects of such relationships, and by expanding the analytical focus from originators (e.g., speakers and authors) to receivers (e.g., listeners and readers).

The *contextual* function refers to the actual space and time of communication and the extent to which they are shared, the communicative purposes, and the perception of the scene (e.g., informal conversation vs. formal written discourse). The *processing* function relates to production and comprehension demands of the communicative event (e.g., writers' clear understanding of their target readers' knowledge base in order to decide what to include or omit in their work). Last but not least, the *aesthetic* function is concerned with the personal or cultural attitudes about the preferred linguistic forms, the ones prescribed by language experts and policymakers, or consonant with individual notions about "good" style and rhetorical influence.

More importantly, this model not only embraces the multi-functionality theory of language, but also proposes the notions that a group of linguistic features can co-occur to serve the same function, that discourses are systematically associated with each other by their deployment of those functions, and that the lexico-grammatical features of a linguistic construction in discourse can be interpreted by determining the functions most frequently and widely served by them.

To conclude, this dissertation will synthesize the functional models of language in this line of research from the semantic, discursive, and systemic functional perspectives. More

DOI:10.6814/NCCU202100169

importantly, this dissertation will follow the long-established tradition in this line of research since the 1960s by looking at the multi-functionality of language. This will be done through examining not only how the target structure is used to simultaneously realize several different functions, but also how different lexico-grammatical features of this structure combine to serve the same function. Finally, by determining the most frequently and widely used functions performed by such features, a comprehensive picture of the use of post-modification in the *Noun Complement* construction can be pieced together at the end of this dissertation.

2.5.4 Analytical framework for description of the post-modification functions in *Noun Complement* construction

Based on the above theoretical frameworks and analytical methods, an analytical framework for description of the target structure (i.e., post-modification in the *Noun Complement* construction) was developed to analyze the functions served by its lexico-grammatical features with significant inferential statistical differences between disciplinary fields and writers of contrasting linguistic-cultural backgrounds.

Specifically, each function of the target lexico-grammatical feature will be first subsumed under the two major types of the functional categorization: (a) those intrinsic and fundamental to language itself (language intrinsic functions), and (b) those outside of language itself (extra-linguistic functions), based on whether it is oriented to the first-order or second-order reality explained in Halliday's systemic functional theory. In terms of language intrinsic functions, the subdivision of the target features will be carried out according to their use (a) to convey informational content or referential meaning (based on Biber and his associates' *informational purpose*, Leech's *informational function*, Lyons's *descriptive function*, Halliday's *ideational metafunction* and Biber's *ideational function*), (b) to provide the complete meaning of the head noun together with the complement clause (based on the functions of general nominal postmodifiers in Biber and his associates' typological theories), and (c) to build links with language itself through marking

DOI:10.6814/NCCU202100169

information structure or cohesion (based on Leech's *aesthetic function*, Halliday's *textual metafunction* and Biber's *textual function*).

The extra-linguistic functions of the target features will be identified by examining in all their coded instances across the subcorpora (a) writers' expression of feelings and attitudes towards content, their purposes of communication, and their perception of the genre of research articles (based on the extra-linguistic functions of pre-modification use in the *Noun Complement* construction, Leech's *expressive function*, Lyons's *socio-expressive function* and Biber's *personal function*), (b) the relationships between discourse participants in research articles, with the writer and the reader end equally attended to (based on Leech's *directive* and *phatic function*, Halliday's *interpersonal metafunction* and Biber's *interpersonal function*), (c) how the production and comprehension demands of research article writing are satisfied (based on the functions of general nominal postmodifiers and Biber's *processing function*), and (d) the cultural preferences towards the forms of language (based on Leech and Biber's *aesthetic function*).

DOI:10.6814/NCCU202100169

CHAPTER 3

Methodology

This dissertation employs a corpus-based and mixed-methods approach to investigate the frequencies, forms, functions and motivations of the post-modification use in the *Noun complement* construction. The quantitative analyses were conducted based on the self-constructed corpus of published research articles across disciplinary and cultural contexts. The inferential statistical methods were applied to identify variances in lexico-grammatical features of the target structure. With the support of the quantitative results, an analytical framework for description of the target structure functions was developed to qualitatively analyze the differences in the functions served by these statistically significant lexico-grammatical features. The statistical results also contributed to constructing the explanatory model of the underlying motivations behind the different use of this structure for qualitative reason analyses.

Methods

3.1 Introduction

The first two chapters have introduced the need for more research into cross-disciplinary and cross-cultural differences in the use of post-modification in *Noun Complement* construction. With the aim of filling this gap, this study needs a corpus that satisfies specific criteria of representativeness, including the number, length, subject and source of texts. Given that no existing corpus constructed using these criteria was available

		RAs	
		Number of texts	Number of words
CIJA	Natural sciences	150	977,389
	Social sciences	150	2,745,058
CCLJA	Natural sciences	150	725,138
	Social sciences	150	1,119,960
Totals		600	5,567,545

Table 3.1 Word counts and number of texts across corpora

DOI:10.6814/NCCU202100169

to the researcher, it was necessary to design and build a new one to accomplish the objectives of the present study. The rest of this chapter is thus to give a detailed description of the methodology adopted to build and analyze this new corpus.

3.2 Corpus description and text selection

This research draws on two self-built corpora, namely, the corpus of international journal articles (CIJA) and the corpus of Taiwanese/Chinese local journal articles (CCLJA) (see Table 3.1).

3.2.1 The CIJA Corpus

The CIJA corpus consists of 300 research articles (RAs) from five international journals in each branch of the natural (chemistry, physics, biology) and social sciences (applied linguistics, law, economics), amounting to 3,722,447 words (for the listing of journals, see Appendix). These six branches were selected because (a) each of them is a major and representative branch in either the natural or social science fields, which, thus, exhibits the defining properties of the corresponding field (Becher, 1994; Becher & Trowler, 2001; Biber & Gray, 2016), (b) six of them together are sufficient to cover a range of academic areas in the natural/social dimension and therefore capable of comprehensively exhibiting the disciplinary variation between these two knowledge domains (Gray, 2011; Hyland & Tse, 2005a), and (c) they are frequently analyzed in corpus-based contrastive research on academic writing (see Egbert, 2014, for a full review). This corpus was then divided into two subcorpora: the international subcorpora of the natural and social sciences.

Selection of journals

All journals included in the CIJA corpus are based in the UK or US and achieve the top-tier ranking in their subject categories based on the five-year impact factor in the 2020 edition of the *Journal Citation Reports* (JCR) (Thomson Reuters, 2019). The reports were made reference to because they are able to provide the valuable and reliable benchmark against which the research community can evaluate the world's high-quality academic journals in a wide variety of research fields. The publishing of work in JCR-

DOI:10.6814/NCCU202100169

indexed journals of the top-tier ranking is usually considered as one of the individual's qualifications for promotion up the university status ladder.

Selection of research article authors

In each journal, 10 original RAs were selected. In this regard, the inclusion criteria established by Wood (2001) and Pan, Reppen and Biber (2016) were used to ascertain a writer's first language. Specifically, native English writers are operationally defined in this study as any author who has a first and last name that can be considered native to English-speaking countries (with the exception of the US writers with non-English surnames given this country whose number of immigrants is highest in the world) and is concurrently affiliated with an institution in a country where English is spoken as the first language. As such, the criteria excluded the authors whose names indicate that they are not native speakers of English (e.g., *Wei Tan*, *Sukhwinder Thandi*, *Mei-Lin Ah-See*), despite being at the institutions located in English-speaking countries. Given that a large number of writers from the US have non-English surnames, only those whose first names appear to be English (e.g., *Dylan* Rodríguez) and affiliated with the institutions in the US were preselected. Before they were recognized as native English writers, their curriculum vitae would be checked to ensure that they had undergone their undergraduate to doctoral education at a university in the English-speaking country. This was done because there is a general consensus in the field of SLA that L2 learners with the age of arrival under 18 to the host country are most likely to reach a native-like ultimate language attainment (e.g., DeKeyser, 2000; MacSwan & Pray, 2005). Therefore, these writers with non-English surnames were at least native-like speakers of English. Moreover, as is standard for natural scientific writing, the vast majority of its research articles are multi-authored. Under these circumstances, the author contributions section in the research article was checked to evaluate the authors who contribute to the writing of the manuscript by the criteria. Only when all these contributors satisfied these criteria, was the article able to be included in the corpus. In contrast, only single- or co-authored articles were taken into account in the field of social sciences, where the trends of multiple authorship have not been prevailing. In the

DOI:10.6814/NCCU202100169

cases of co-authored work, both authors needed to meet the above criteria. To sum up, this study erred on the side of excluding authors with questionable names.

Identification of article numbers per journal

Extracting the exact number of 10 RAs from each journal was under inferential statistical considerations. Specifically, this enabled one subcorpora to be presented by 150 RAs (sample size, n = 150), which is, thus, sufficient enough to run the nonparametric Mann-Whitney (MW) U test used in the present study (still applied even with small sample sizes, n = 40-60) (for details about statistical analyses, see Section 3.5). More importantly, Wang et al. (2003) devised the optimal sample size formula for the MW test

$$n = \frac{(Z_{\alpha/2}/\sqrt{12} + Z_{\beta}\sqrt{p_3 + 4p_4 - 4p_2^2})^2}{(1/4 - p_2)^2},$$

where α and β denote the type-I and type-II error rates respectively, the Z_is are generated from normal population with mean y and variance 1, and the p_is are estimated by Monte Carlo method based on a sample of size 10,000 (i = 1, ... , n). Following this, the average 9.6 RAs (for eight tests, rounded off to 10) per journal were needed in order to have an 80% power to confirm the observed difference between two subcorpora when such a difference truly exists. In other words, the optimal sample size of the MW test in this study entailed the selection of 10 RAs from each journal. Review articles, editorials, commentaries, letters to the editor, etc. were all excluded, since they do not specifically reflect the genre of research article writing that makes original contributions to the field.

3.2.2 The CCLJA Corpus

The CCLJA corpus (totaling 1,845,098 words) was similarly compiled through the selection of 300 RAs of the same six branches. From each of them, five local journals of the upper rank in citation index were chosen, as per the *Report on the International and Domestic Citation of China's English Academic Journals* (CNKI-EJCR, 2017). The decision to choose this report was made because it offers Taiwan and mainland China's most influential ratings on their domestic English journals, and thus that it presents a

DOI:10.6814/NCCU202100169

representative benchmark of local writers with the high levels of research and English language expertise. The methods adopted to ascertain the writer's first language as Chinese were relatively simple, given the local journals whose authorship belongs mostly to Chinese native speakers. Chinese EFL writers were operationally defined as any author affiliated with an institution in either Taiwan or mainland China, while also having a first and last name considered native to Taiwan or mainland China. In the cases of multi-authored articles in both the natural and social sciences, only the ones whose joint authorship all met the above criteria were included in the corpus. This corpus was also divided into two subcorpora, namely, the local subcorpora of the natural and social sciences. All collected texts in the corpus were published after 2015.

3.3 Text formatting and cleaning procedures

When the collection of the 600 corpus files was completed, each text in its original PDF format was edited manually through the professional PDF file editor *Adobe Acrobat Pro DC* (Adobe Inc., 2020). This was done because the subsequent conversion to plain text was problematic due to the page layout and design of the specific journal. Specifically, all page headers and footers (which consist of the journal, article or author information and a page number), visual aids (i.e., pictures, tables, figures, charts, graphs), and reference lists were removed from the files, "as they are not a part of the language of the article itself" (Gray, 2011: p. 51). On the other hand, captions of all visual aids and footnotes were retained, granted that they tend to include plenty of material pertaining to background information, claims and arguments, and even data (see, e.g., Conrad, 1996). However, the footnotes where references were cited in full (particularly in law journal articles) were deleted, while those appearing at the bottom of the pages were relocated to the end of the file. The RAs with the two-column layout (mainly in natural scientific journals) were switched to the single-column ones. This, along with the relocation of footnotes, contributes to the accuracy of file conversion, for the existing converter software is not able to distinguish the footnotes from the main body of the text, while also treating all column layouts as the one-column format. Without these two cleaning processes, the data

DOI:10.6814/NCCU202100169

output would turn out to be so scrambled that the automatic tagging errors were most likely to increase exponentially in the succeeding annotation procedure.

Additionally, texts included in natural scientific subcorpora contain considerable formulas and special symbols, which were, once set apart from the body of the text on their own lines, deleted, whereas they were kept, if inserted into the text of the article. All other aspects of the source text, including the use of identification, bold, italics, headings and sub-headings, were retained, as they form part of the article language itself.

Following all these editing processes, each file was converted into plain text using the converting software *PDF Converter Elite* (PCE, 2017). A standardized header was then added to the beginning of each plain text file for recording the bibliographic information on the RA. All text files were also given a descriptive filename that contains the following information: branch, nationality (international, local), a unique identification number, and source.

3.4 Tagging scheme and corpus annotation

To ensure the efficiency of extracting *Noun Complement* constructions from huge quantities of data, a part-of-speech (POS) tagger scheme was applied to tagging all the running words in the corpus with their parts of speech (e.g., noun, determiner, adjective, etc.). POS tagging was conducted through *TagAnt* (v. 1.2.0) (Anthony, 2015), built on the *TreeTagger* engine (Schmid, 1995) whose trigram version (adopted in the present study) can achieve 96.36% accuracy on the quality of the tagging result (Schmid, 2003). To check its accuracy, the words in five RAs randomly selected from each corpus (CIJA and CCLJA) were manually parsed. A comparison between the automatic and manual parsing of the RAs shows that the *TagAnt* program achieved an accuracy rate of 96% and 94% respectively. The lower accuracy rate for the RAs in local journals results mainly from the program's failure to identify the Chinese characters inserted into the body of the sentence (e.g., the original Chinese title of a legal document or economic report). Notwithstanding this, the accuracy levels were acceptable, so the frequency normalization was based on the item counts yielded by the *TagAnt* program.

DOI:10.6814/NCCU202100169

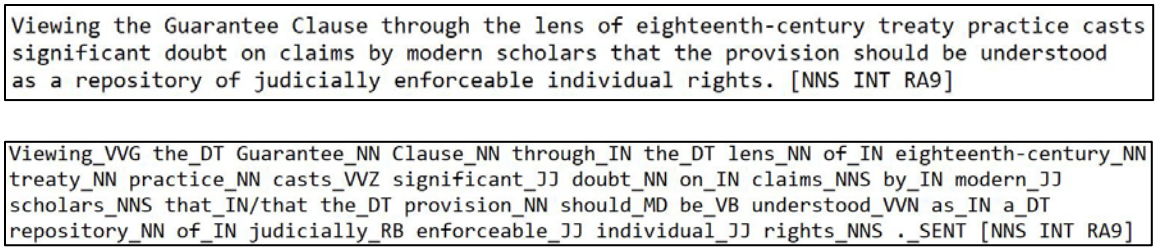


Figure 3.1 Examples of original (above) and POS-tagged (below) text via *TagAnt*

Granted that in the *TreeTagger* tagset, the word *that* is particularly tagged with the part of speech *complementizer* (besides its other word classes *determiner*, *pronoun*, *conjunction* and *adverb*), the cases where *that* is not used to subcategorize a complement clause were automatically excluded (e.g., relative clause). In contrast, it must also be acknowledged that extracting the items with the omission of *that* in the *Noun Complement* construction is technically impossible through the *TreeTagger* system. However, the presence of *that* prior to a nominal complement clause has traditionally been taken as the norm in academic writing (e.g., Biber et al., 1999; Stubbs, 1996; Swan, 2005). Even in an informal style, "[*t*]*hat* is not usually dropped after nouns" (Swan, 2005: p. 265), as illustrated in the example below.

He disagreed with Copernicus' view that the earth went round the sun.

(NOT . . . ~~*Copernicus' view the earth went*~~ . . .) (ibid.)

Thus, this entails the extreme rarity of the *that* omission in RAs. Also, a backcheck on three most frequently used nouns in the present corpus, namely *fact*, *evidence*, and *belief*, confirmed this usage pattern (*that* not dropped). As a result, only the instances where *that* is retained were included in the analysis.

After the POS-tagging procedure, the tagged texts were manually checked again to secure the further improvement of the tagging accuracy. Figure 3.1 compares one example of the POS-tagged text with the original text.

In what follows, the concordance software *AntConc* (Anthony, 2019) was run to search for the *Noun Complement* construction on the basis of syntactic information through regular expression queries:

DOI:10.6814/NCCU202100169

1. For the search of *that*-clauses:

   *\w*_[^VJ]\w*\s*that_IN/that*

2. For the search of *to*-clauses:

   *\w*_[^VJ]\w*\s*to_TO\w*\s*\w*_V*

3. For the search of *of + ing*-clauses:

   of_IN\w\s*\w*_V*

4. For the search of *wh*-interrogative clauses:

   *w*_(WDT|WP|WRB)\w*\s*\w*_[^VM]*

A manual reading of the concordance lines was then carried out to identify the target feature *post-modification* in the *Noun Complement* construction. Specifically, to ensure that only *that*-complement clauses were selected in the data, the running word *that* in the corpus was manually checked to see whether it was tagged with the part of speech *complementizer* correctly (POS tag: *IN/that*), in a way that completely excluded any demonstrative and relative pronouns of *that*. Following this, the target feature for analysis was located and then coded for each item according to its grammatical category. In terms of *to-infinitive* and *of + ing*-complement clauses, the same coding procedure was conducted to pinpoint the target constructions before classifying them based on their linguistic forms. Finally, after the extraction of all *wh*-interrogative clauses in the tagged data, each *wh*-word in the corresponding clause was checked to make sure its word class as the *complementizer*. Both delimiter and line-breaking functions within the tool preferences of *AntConc* were used to arrange the query term (i.e., *that, to, of, wh-word*) in its own space as a way to ease interpretation of the output for analysis.

Identification of the *Noun complement* construction

Given that the focus of this study is on the use of post-modification in the *Noun Complement* construction where a stance head noun takes a nominal complement clause, each head noun of the post-modification structure was checked to ensure that it belongs to the categories of stance nouns in Hyland and his associates' (e.g., Jiang, 2017; Jiang &

DOI:10.6814/NCCU202100169

Hyland, 2015) classification of stance nouns in the *Noun Complement* construction (see Table 3.2).

In their classification, stance nouns are functionally used to mark ***entities***, describe ***attributes*** of entities, and discuss ***relations***. Nouns included in the category of ***entity*** orient to objects, events, discourses, and aspects of cognition. Specifically, ***object*** nouns pertain to concrete metatext, and hence *report, paper* and *extract* are typical examples in this subcategory. ***Event*** nouns relate to actions, processes or state of affairs which have a spatiotemporal location, thus containing nouns such as *change*, *process* and *evidence*. ***Discourse*** nouns project a stance on verbal propositions and speech acts, such as *argument, claim* and *conclusion*. Nouns which concern ***cognition*** refer to beliefs, attitudes and elements of mental reasoning, with nouns such as *decision*, *idea*, *belief* and *doubt* included in this subcategory.

As to the category of ***attribute,*** it concerns the evaluations of the quality, manner, and status. ***Quality*** nouns are those with traits of value judgement, such as *advantage, difficulty,* and *value*, thereby assessing whether something is praised or criticized, valued or depreciated. ***Manner*** nouns, by contrast, describe the circumstances and formation of

Entity	**Description**	**Examples**
Object	Concretizable metatext	*Report, paper, extract*
Event	Events, processes, states of affairs	*Change, process, evidence*
Discourse	Verbal propositions and speech acts	*Argument, claim, conclusion*
Cognition	Cognitive beliefs and attitudes	*Decision, idea, belief, doubt*
Attribute	**Description**	**Examples**
Quality	Traits that are admired or criticized, valued or depreciated	*Advantage, difficulty, value*
Manner	Circumstances of actions and state of affairs	*Time, method, way, extent*
Status	Epistemic, deontic and dynamic modality	*Possibility, trend, choice, ability*
Relation	**Description**	**Examples**
Cause-effect, difference, etc.	Cause-effect, difference, relevance	*Reason, result, difference*

Table 3.2 Classification of stance nouns in the *Noun Complement* construction

DOI:10.6814/NCCU202100169

actions and states of affairs, which involve nouns such as *time*, *method*, *way* and *extent*. Nouns which relate to ***status*** make judgement of epistemic, deontic and dynamic modality. Epistemic modality is concerned with possibility and certainty, such as *likelihood* and *truth*; deontic modality denotes obligation and necessity, such as *need* and *obligation*; dynamic modality depicts ability, opportunity and tendency, such as *authority*, *potential* and *tendency*. Finally, stance nouns can be employed to make a stance by specifying how a writer comprehends the connection or relationship to the content in a proposition, expressing ***relations*** such as *reason*, *result* and *difference*.

Based on this categorization, each head noun of the extracted post-modification structures were examined and manually selected by referring to the specific definitions and examples of the above stance noun categories. Finally, there were totally 5,110 head nouns of post-modification identified as stance nouns and thus included in the analyses of the present study.

The following categorization of the target feature was carried out by coding each instance of the extracted post-modification structure based on its lexico-grammatical category (e.g., phrases, clauses, etc.). Table 3.3 illustrates the eight lexico-grammatical features of the post-modification use extracted and identified through the annotation and categorization procedure:

Categories	Subcategories	Examples
Phrases	Prepositional phrase	Thereby we countered the long-held belief **in aqueous remediation literature** [that an increase in surface area results in an increase in sorption capacity].
	Adverb phrase	There is clearly a need for both an objective evidence base for its application (hence our suggestion **above** [of applying economic data and data on the demand for public services]).
	Adjective phrase	Interestingly, the authors reached a conclusion **similar to ours** [that both methods measure essentially the same phenomenon and that the bias can be explained by the different assumptions that the tests are based on].

DOI:10.6814/NCCU202100169

	Noun phrase	If 16 known carrier mares could be located for breeding, failure of those 16 matings to produce a CID foal would provide statistical assurance **(1% level)** [that the stallion was not a CID carrier].
Clauses	Non-finite clauses	In this case, the national authorities were free to provide for an appeal with suspensive effect against decisions **relating to an EAW** [to uphold domestic constitutional rights], because the EAW Framework Decision left them a degree of discretion in this regard.[105] This contrasts with previous population studies, which have shown no gender differences in the prevalence rate of BPD (e.g. [4, 5, 10]); yet it confirms the suggestion **made in the DSM-5** [that BPD is more common among women [30]] and also substantiates the findings of clinical studies, which have demonstrated that more women than men suffer from BPD. Yet the model does not overdetermine how reality is mapped to vocabulary: there is often a choice **to be made** [in how entities are to be described and which perceptual or interactional properties of an entity are highlighted].
	Finite dependent clauses	Then the implicit assumption **which seems to be used** [that advocated and enacted policies are identical] might be more palatable. This important effect of malaria on LBW is confirmed by the significant trend **we found** [of increasing risk of LBW with repeated infections of the mother over the entire period of pregnancy].
Punctuation		The nature of the relationship reflects the overarching hypothesis**,** [that increased exposure will lead to increased expectation (and therefore perception), with the exception of salient boundary zones, occupied here by the San Diegans]. To get both employment and output deindustrialization, we need to make additional assumptions**:** [that the trade balance in manufactures becomes more negative or that there is a secular demand shift away from manufactures].
Extraposition		But the simple fact **remains** [that narrative mosaics cannot really be assessed in any other way].

DOI:10.6814/NCCU202100169

	However, because time and space are conveniently interchangeable at ultrafast shutter speeds (time range less than 1 ns), the question **arises** [of whether the design of imaging optics for an ultrafast camera should follow the same rules as for conventional low-speed cameras].

Table 3.3 Lexico-grammatical features of post-modification use in the *Noun Complement* construction

3.5 Statistical analyses

To draw quantitative comparisons, the frequency of the target feature *post-modification* was counted, along with each of its lexico-grammatical feature. Given the different sizes of the subcorpora, the raw frequencies of the coded items for each file were normalized by 10,000 words for all RAs, thus minimizing any effect caused by varying corpora sizes (see, e.g., Biber, 2006a). After this, these standardized frequencies were exported on a file-by-file basis to an Excel spreadsheet, which was, then, imported into the *IBM SPSS* statistics program (v. 24.0) (IBM Corp., 2018) to ascertain the significance of variation within the examined items across the subcorpora. Specifically, granted that the corpus data at many of the data points were determined to be non-normally distributed (following significant Shapiro-Wilks p values: $p < .001$), the non-parametric Mann-Whitney *U* comparisons, with Holm–Bonferroni correction on the *p* values, were employed to determine the variance between variables across subcorpora. An alpha value was set at 0.00625 (for 8 tests) to correct for multiple testing (avoiding "data dredging"/"data fishing", where multiple tests can lead to false significance, see Sowey & Petocz, 2017: p. 129-130). Effect sizes (Pearson's *r*) were calculated through the following formula:

$$r = \frac{Z}{\sqrt{N}},$$

where *N* is the total number of cases, with the *Z* score obtained from the Mann-Whitney *U* test result produced by the *SPSS* program. The absolute values of the effect size represent the magnitudes of difference, which can be interpreted as: $r = 0.1$ (small effect), $r = 0.3$

DOI:10.6814/NCCU202100169

(medium effect), r = 0.5 (large effect) (Cohen, 1977; Field, 2009). Only the lexico-grammatical features with a Pearson r value of at least ±.300 meet the threshold for inclusion of the present study.

3.6 Qualitative analyses

With the support of quantitative analyses, the lexico-grammatical features with significant inferential statistical differences between disciplinary fields and writers of contrasting linguistic-cultural backgrounds (effect size (r) ≧ ±.300) were identified. Qualitative analyses were then conducted with the aim of providing comprehensive accounts of the functions of these statistically different lexico-grammatical features and the underlying motivations for the use of the target structure. Firstly, in terms of functional analysis, the analytical framework for description of the post-modification functions in *Noun Complement* construction (see Section 2.5.4 for details about this framework) was employed to analyze the differences in the functions served by these lexico-grammatical features. Specifically, each function of the target feature was first categorized into two major types – language intrinsic functions and extra-linguistic functions. The functions included in the language intrinsic category were then subdivided into three further groups, based on their use to (a) convey informational content or referential meaning, (b) provide the full meaning of the head noun together with the complement, or (c) build links with language itself through marking information structure or cohesion. The extra-linguistic category, in contrast, was subcategorized into four groups, according to the roles the target features play in all their coded instances across the subcorpora that concern (a) the writer (whose expression of feelings and attitudes towards content, purposes of communication, and perception of the genre of research articles), (b) the relationship between writer and reader, (c) the production and comprehension demands of research article writing (genre-specific demands; e.g., writers' decision about what to include in their work through the prediction on target readers' knowledge base), and (d) the cultural preferences (for details about the theoretical backgrounds of this analytical framework, see Chapter 2, Section 2.5). Moreover, based on the multi-functionality theory of language, the functional analysis also

DOI:10.6814/NCCU202100169

examined how the target structure is used to simultaneously realize several different functions, and how different lexico-grammatical features combine to serve the same function.

The explanatory model of cross-disciplinary and cross-cultural differences in the use of post-modification in the *Noun Complement* construction was developed to reveal the underlying motivations and enabling factors behind the different use of the target structure. As noted in Chapter 2, studies into the three major components of the *Noun Complement* construction (i.e., stance noun, proposition in the complement clause, and pre-modification) in academic writing report the reasons that underlie the differences in disciplinary use of these components in terms of their frequencies, forms and functions. These reasons were exclusively attributed to the disciplinary-specific factors which center around the modes of knowledge construction that point to value systems and ideologies of the specific academic fields. Although these analyses were not conducted within a wider social context where the actual research work occurs, they still guided the construction of this explanatory model to include a rhetorical dimension, while also helping to treat disciplinary modes of knowledge construction as a factor in this dimension (see Table 3.4 below). Similarly, disciplinary rhetorical traditions, such as less hedging in hard sciences, pertain directly to the rhetorical dimension and were thus added to it as well.

Moreover, post-modification in the *Noun Complement* construction is itself a linguistic structure, with the reasons that underlie the differences in disciplinary and cultural use of its lexico-grammatical features and functions explored by this explanatory model. Therefore, a linguistic dimension is naturally part of this model. In this regard, linguistic conventions help to generate the structure, content and context of scientific research writing (Hyland, 1998c), where the way a discipline exploits grammatical patterns and lexical resources and the extent to which it privileges certain ones over others is determined by the epistemology and ideology of the discipline concerned (Charles, 2006). As such, linguistic conventions in the genre of research articles are also disciplinary-specific, which is, in turn, included as a factor in the linguistic dimension. The disciplinary differences in

DOI:10.6814/NCCU202100169

Cross-disciplinary differences			
Linguistic dimension	**Rhetorical dimension**	**Relational dimension**	**Socio-cultural dimension**
Linguistic conventions in the genre of research articles	Modes of knowledge construction	Interpersonal writer-reader relationship	Developmental differences in scientific research
Nature of the grammatical complexity differences	Disciplinary rhetorical traditions		Social and political movements in Western cultures
Cross-cultural differences			
Linguistic dimension	**Rhetorical dimension**	**Relational dimension**	**Socio-cultural dimension**
Semantic, syntactic and punctuation systems in the Chinese language	Traditions of Chinese rhetoric	Writer-responsibility versus reader-responsibility	Chinese cultural practices rooted in Confucian and Taoist traditions
Chinese EFL writers' language competence in English academic writing	Chinese EFL writers' academic competence in English academic writing		Social and political movements in Western cultures

Table 3.4 Explanatory model of cross-disciplinary and cross-cultural differences in the use of post-modification in the *Noun Complement* construction

grammatical complexity illustrated in Biber and his associates' typological theories noted above also inspired the present study to add the nature of the grammatical complexity differences as another factor in the linguistic dimension.

Additionally, according to Ädel (2006) and Hyland's theories (1999a, 2005b, 2007) about academic metadiscourse, metadiscourse is defined as the aspect of a text which explicitly organizes the writer's stance-making. In this respect, the interpersonal relationship between writer and reader is the major concern of metadiscourse. Illustratively, Figure 3.2 below shows that a writer is able to make a stance in the two "worlds" (the world of discourse and the "real world"). However, when the stance is projected in the text, the complete overlap of metadiscourse and stance-making occurs (i.e., metadiscourse equals

DOI:10.6814/NCCU202100169

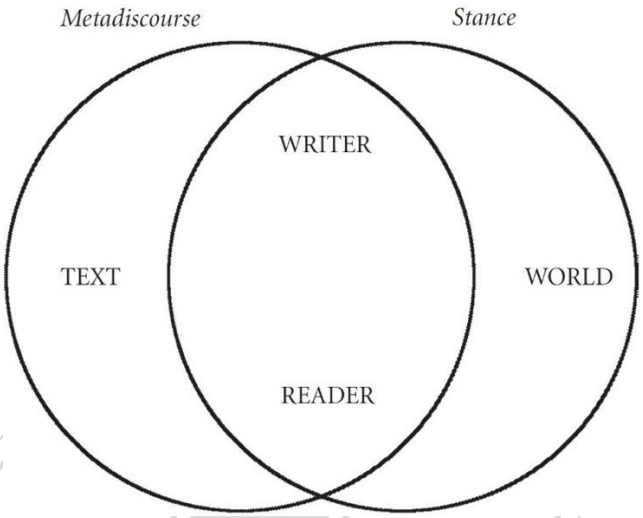


Figure 3.2 Components and ranges of metadiscourse and stance (Ädel, 2006: p. 41)

stance), which, then, has at its core the interpersonal relationship between writer and reader. There is no doubt that post-modification, as a stance-making construction, participates in organizing the writer's projection of a stance in the world of discourse, and thus attend particularly to the interpersonal writer-reader relationship as well. Hyland (1999a, 2005b, 2007) also notes the disciplinary differences in dealing with the interpersonal writer-reader relationship. That is, writers in soft fields place high importance on establishing and maintaining the interpersonal relationship with readers, given the discursive nature of the lines of inquiry in soft knowledge domains, thus producing reader-inclusive texts. On the other hand, hard field writers lay great stress on "the impartiality and linearity of science production" (Hyland & Jiang, 2016: p. 32). As such, the relational dimension was added to the explanatory model, which, in turn, has the interpersonal writer-reader relationship as its core.

The social-cultural dimension was also incorporated into this model, as writing in any specific discipline is situated in a wider social world (Hyland, 2004), and the differences in cultural use of the target structure are central to this study. This dimension is mainly connected with the wider world where the research work actually takes place. It thus

DOI:10.6814/NCCU202100169

concerns the general development of scientific research in the modern socio-economic world. Actually, given the globally competitive environment for science and technological innovation, scientific research is becoming increasingly complex even on a daily basis mainly in hard fields. In this way, information explosion, for example, is more likely to happen in hard fields, therefore leading to the economy of expression, which, then, causes the differences in disciplinary use of post-modification structures. Such developmental differences in scientific research are also reflected in the more interdisciplinary research trend in the natural than social sciences, according to van Noorden's (2015) meta-analysis of 35 million papers from varying disciplinaries published from mid-1980s to 2015. This, for example, may change the composition of the readership in the hard sciences (larger proportion of less specialized audiences) and how natural scientists handle the interpersonal writer-reader relationship.

Social and political movements in Western cultures (e.g., post-modern and feminist movements) is interpreted as a factor too, since these movements are directly attached to the soft fields of *philosophy*, *sociology*, *literary criticism*, etc. (see, e.g., Ashenden, 1997; Garry & Pearsall, 2015; Vidal, 1991), thereby exerting a more profound influence on social scientists than their natural counterparts, and leading to disciplinary differences in the use of post-modification.

The dimensions and their respective factors of the explanatory model in terms of cross-cultural differences were identified based on the factors of English competence and L1 influence revealed in the prior research into cultural differences in the use of the *Noun Complement* construction. These two factors actually helped to construct the linguistic dimension with the focus on the idiosyncratic features of the semantic, syntactic and punctuation systems in the Chinese language, as well as Chinese EFL students' levels of proficiency in English. The traditions of Chinese rhetoric and Chinese EFL students' academic competence (which mainly concerns their familiarity to English academic register and genre conventions) were made central in the rhetorical dimension. The relational dimension concerns the interpersonal writer-reader relationship in terms of

DOI:10.6814/NCCU202100169

writer-responsibility in the West versus reader-responsibility in the East. The former holds writers primarily responsible for effective communication, while the latter places this responsibility on readers. The socio-cultural dimension includes the factors of the Chinese cultural practices rooted in Confucian and Taoist traditions, and also the influence from the social and political movements in Western cultures.

Under this model, the reasons that underlie the cross-cultural differences were first analyzed. Specifically, in the linguistic dimension, the reasons were attributed to factors relating to the idiosyncratic features of the semantic, syntactic and punctuation systems in the Chinese language, and Chinese EFL students' English competence. The traditions of Chinese rhetoric and Chinese EFL students' academic competence were taken into accounts in the rhetorical dimension to pinpoint the reasons. As to the relational dimension, the analysis of the reasons was conducted depending on the factor of interpersonal writer-reader relationship, by comparing the notions of writer-responsibility in the West and reader-responsibility in the East. The socio-cultural dimension was responsible for providing the reasons from the perspectives of the Chinese cultural practices rooted in Confucian and Taoist traditions and the influence from the social and political movements in Western cultures.

In the similar vein, the reasons that underlie cross-disciplinary variations were explored by considering the factors of linguistic conventions and grammatical complexity differences in the linguistic dimension, of knowledge construction modes and rhetorical traditions in the rhetorical dimension, of writer-reader relationship (reader-inclusive or reader-exclusive) in the relational dimension, and of developmental differences in research and socio-political movements in the socio-cultural dimension.

It should be noted here that since the aims of the present study are to examine the differences in disciplinary and cultural use of post-modification in the *Noun Complement* construction, only the lexico-grammatical features with significant inferential statistical differences between disciplinary fields and writers of distinctive linguistic-cultural backgrounds were included into the qualitative analyses. Similarly, the reasons analyzed

DOI:10.6814/NCCU202100169

in this study were narrowed down to why the significant inferential statistical differences exist in disciplinary and cultural use of the target structure, not the general reasons why the academic writers use this structure, which, to a large extent, restricts the reason analysis to less dimensions of the explanatory model. Therefore, although each statistically different lexico-grammatical feature and its corresponding functions have been attempted to be analyzed from all four dimensions of the explanatory model, the underlying reasons of disciplinary and cultural differences were attributed to the varying factors in a number of rather than all dimensions.

Inter-rater reliability

For the categorization of the lexico-grammatical features of the target structure, and the contruction of the analytical framework for functional analysis and the explanatory model of cross-disciplinary and cross-cultural differences, two faculty members at the universities in Taiwan and the US with research and teaching experience of over ten years in the field of applied linguistics and L2 writing were invited to analyze the samples. The inter-rater correlation (Pearson's *r* in the *SPSS* program) for the categorization is 0.96; the analytical framework, 0.91; and the explanatory model, 0.93, which, therefore, indicates a high degree of agreement.

DOI:10.6814/NCCU202100169

CHAPTER 4

Results

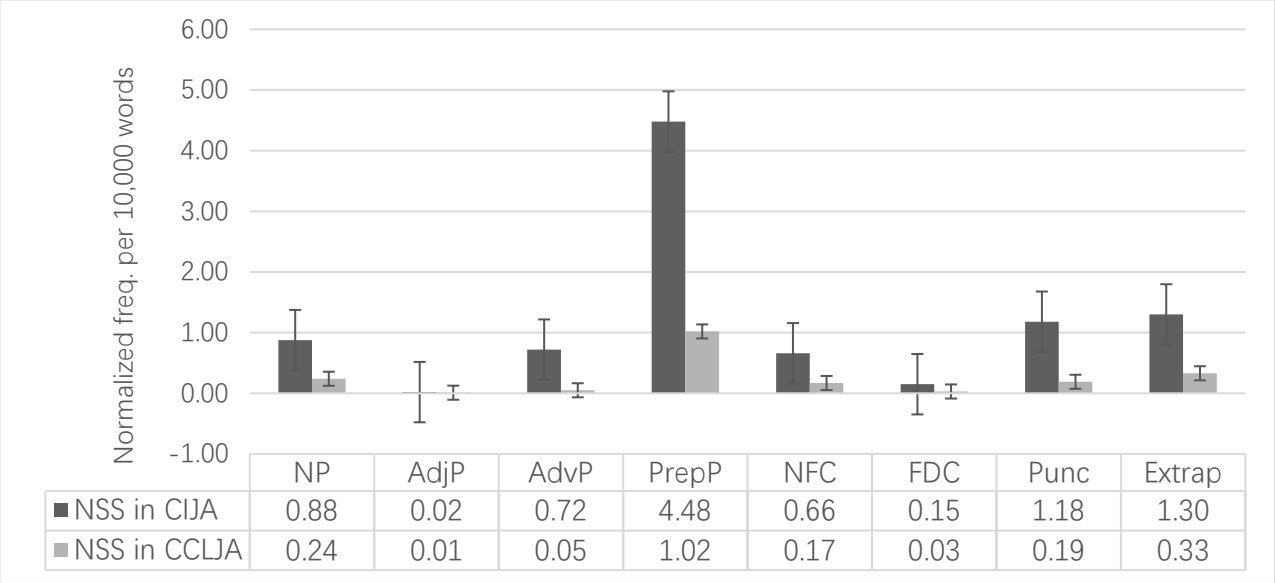
4.1 Introduction

This chapter first presents the inferential statistical results of the variances in the lexico-grammatical features of post-modification in the *Noun Complement* construction. The following qualitative analyses then drew on the analytical framework for description of post-modification functions to generate the results of the cross-disciplinary and cross-cultural differences in the functions served by the lexico-grammatical features with statistical significance (effect size (r) ≧ ±.300). Both statistical and qualitative results are presented in order of the comparisons between (a) international and local writers in the natural sciences, (b) international and local writers in the social sciences, (c) international writers in the natural and social sciences, and (d) local writers in the natural and social sciences.

4.2 International vs. local writers in the natural sciences

Figure 4.1 below compares the frequency of post-modification use in terms of its lexico-grammatical features by international and local writers in the natural sciences. Overall, the lexico-grammatical feature *prepositional phrase* is most preferred by both international and local writers in their post-modification use (4.48 vs. 1.02), while they all make least use of *adjective phrase* as post-modification, with its average normalized (per 10,000 words) values across the two contrasting subcorpora plummeted to 0.02 and 0.01 respectively. Specifically, the lexico-grammatical feature *extraposition* is international natural scientists' second preference, which is followed by *punctuation*, *noun phrase*, *adverb phrase*, *non-finite clause,* and *finite dependent clause*. Local natural scientists, on the other hand, also tend to favor *extraposition* secondly, and then, with some slight differences from their international counterparts, *noun phrase*, *punctuation, non-finite clause, adverb phrase*, and *finite dependent clause*. Moreover, significant inferential statistical differences were found in post-modification use between international and local natural science writers (see Table 4.1 below). In general, the difference pertaining to the

DOI:10.6814/NCCU202100169



	NP	AdjP	AdvP	PrepP	NFC	FDC	Punc	Extrap
NSS in CIJA	0.88	0.02	0.72	4.48	0.66	0.15	1.18	1.30
NSS in CCLJA	0.24	0.01	0.05	1.02	0.17	0.03	0.19	0.33

Note: NP = Noun phrase, AdjP = Adjective phrase, AdvP = Adverb phrase, PrepP = Prepositional phrase, NFC = Non-finite clause, FDC = Finite dependent clause, Punc = Punctuation, Extrap = Extraposition.

Figure 4.1 Lexico-grammatical features of post-modification use between the Natural Scientific Subcorpora (NSSs) in CIJA and CCLJA (normalized frequency per 10,000 words)

Feature	Mann-Whitney			Effect size (*r*)
	U	Z	*p*	
Noun Phrase	**9312.5**	**-4.041**	**< .001**	***r* = -0.233**
Adjective Phrase	11249.5	-0.005	= .996	*r* = -0.000
Adverb Phrase	**8328.0**	**-6.261**	**< .001**	***r* = -0.362**
Prepositional Phrase	**5107.0**	**-8.558**	**< .001**	***r* = -0.494**
Non-finite clause	**9311.0**	**-4.278**	**< .001**	***r* = -0.247**
Finite dependent clause	**10282.0**	**-3.412**	**= .001**	***r* = -0.197**
Punctuation	**7506.5**	**-6.586**	**< .001**	***r* = -0.380**
Extraposition	**7631.0**	**-6.110**	**< .001**	***r* = -0.353**
Totals	**3380.0**	**-10.675**	**< .001**	***r* = -0.616**

Table 4.1 Lexico-grammatical features of post-modification in the *Noun Complement* construction between NSSs in CIJA and CCLJA

DOI:10.6814/NCCU202100169

total frequency of post-modification use (Mann-Whitney: U = 3380.0, Z = -10.675, p < .001, Effect size: r = -0.616) is clearly discernible given the large effect size reported (|r| > .500). More importantly, this statistical difference is particularly noticeable for the lexico-grammatical feature *prepositional phrase*, considering the almost large effect size found (r = -0.494, p < .001), with medium effect sizes reported for *punctuation* (r = -0.380, p = .001), *adverb phrase* (r = -0.362, p < .001) and *extraposition* (r = -0.353, p < .001), whereas for the rest of the lexico-grammatical features, there exist merely small statistical differences in their use between international and local natural scientists (|r| < .300). Therefore, only the use of *prepositional phrase*, *punctuation*, *adverb phrase* and *extraposition* as post-modification was included in the analyses of the present study, inasmuch as the linguistic features with a Pearson r value of at least ±.300 (medium effect size) are able to meet the threshold for inclusion of this research.

Table 4.2 reports the variations in the functions served by the lexico-grammatical features with statistical significance in international and local natural scientists' post-modification use. Specifically, the most significant inferential statistical difference was found in their use of prepositional phrases (see Table 4.1 above), whose functions are to indicate whether stance nouns are overtly averred by writers, or attributed to other humans or abstract entities, as well as to qualify stance nouns while also functioning as the subjects of the following propositions. To illustrate, the prepositional phrase *of our model* in Example (6) serves to show that the stance noun *assumption* is explicitly averred by the writers of this research article, while in Example (7), the prepositional phrase *by Russell* attributes the stance noun *finding* to another human (i.e., Russell). Similarly, Example (8) shows the attribution of the stance noun *determination* to an abstract entity through the prepositional phrase *by ssNMR spectroscopy*. Additionally, the content of the stance noun *requirement* is, as shown in Example (9), specified in its complement "to cross the impenetrable *Mtb* cell wall barrier", although the two are separated by the prepositional

DOI:10.6814/NCCU202100169

NSSs in CIJA & CCLJA		
Feature	**Function**	**Variance**
Prepositional phrase	Self-sourced	┤
	Other human	+++++*1 ┤
	Abstract entity	−
	Qualifier + subject	+++
Punctuation		
Colon	Complex grammatical structures	+++++*23 ┤
Comma	Identity relationship	−
Dash	Informality	−−−−−*1 −−−−
Adverb phrase	Endophoric marker	+++++*8 ┤
	Code glosses	−
Extraposition	Information principle	+++++
	Principle of end-weight	+++++

Note: + | − : relative risk ratio (*RR*, ratio of two proportions or probabilities) of the corresponding lexico-grammatical feature included in NSS in CIJA to NSS in CCLJA > 1.1 (minus = NSS in CCLJA to NSS in CIJA); ++ | −− : relative risk ratio > 1.3; +++ | −−− : relative risk ratio > 1.5; ++++ | −−−− : relative risk ratio > 1.7; +++++ | −−−−− : relative risk ratio > 1.9; +++++*n | −−−−*n : relative risk ratio = 2^n; ┤|− : relative risk ratio = 0-1.1.

Table 4.2 Lexico-grammatical features, functions and variance of post-modification use between NSSs in CIJA and CCLJA

(6) Despite the success of such a simple model, additional experiments are currently desirable to test the main assumption **of our model** [that both the BDP shift and the saturation effect are due to the influence of the advancing a/c interface on MPDs generated in the crystal bulk and increase of MPD diffusion length with increase of ion fluence due to saturation of defect sinks in the bulk].

[NSS INT RA5]

(7) Consistent with this mechanism is the finding **by Russell** [that photogenerated cyclohexyl radicals undergo reaction with β-phenylethynyl phenyl sulfone (8)[7] to generate phenyl cyclohexyl acetylene (10) and phenyl sulfonyl radical (which does not fragment to sulfur dioxide and phenyl radical,[8] thereby unable to propagate a similar C−H activation event)].

[NSS INT RA11]

DOI:10.6814/NCCU202100169

(8) A second CSP study (CSP2) was commissioned following the discovery of forms VII-IX and subsequent determination **by ssNMR spectroscopy** [that form VII is a $Z' = 2$ polymorph (SI Figure S14)].

[NSS INT RA8]

phrase *for the molecule*. In this regard, the prepositional phrase in question not only qualifies the stance noun *requirement* (i.e., limiting the *requirement* to only the one imposed on a given molecule, thus making it less general), but also functions as the subject of the following proposition "to cross the impenetrable *Mtb* cell wall barrier" (i.e., it is this very molecule that crosses the impenetrable *Mtb* cell wall barrier). As such, this type of functions is named *qualifier + subject*.

(9) This approach has several advantages over current strategies that include: no requirement **for the molecule** [to cross the impenetrable *Mtb* cell wall barrier] and the potential avoidance of drug efflux challenges.

[NSS INT RA4]

Statistically, the *other human* function exhibits the most cross-cultural differences (*RR* = 3.00, large relative effect size/variance), which is, then, followed by the functions of *qualifier + subject* (*RR* = 1.55, small relative effect size/variance), *abstract entity* (*RR* = 1.16, small relative effect size/variance), and *self-sourced* (*RR* = 0.90, small relative effect size/variance).

Moreover, the inferential statistics results show the use of punctuation as post-modification to be the second lexico-grammatical feature that most clearly distinguishes writers with different first languages and cultural backgrounds (see Table 4.1 above). The functions served by this feature include signaling forthcoming complex grammatical structures by colons (*RR* = 46.67, very large relative effect size/variance), marking identity meaning relationships by commas (*RR* = 1.25, small relative effect size/variance), and suggesting informalization by dashes (*RR* = 3.76, large relative effect size/variance medium cross-cultural difference). Example (10) is typical in the use of colons as post-modification:

DOI:10.6814/NCCU202100169

(10) If we considered only the first and second person pronouns/contractions dimension, we would arrive at a quite different set of conclusions**:** [that conversations and panel discussions are quite similar, fiction and academic prose are quite similar, and these two sets of texts are quite different from each other].

[SSS INT RA23]

In this excerpt, the colon is placed after the stance noun *conclusions* to indicate that its complement clauses of considerable grammatical complexity (i.e., the coordinated *that*-clauses) are about to come. Or, in Sanchez-Stockhammer's (2016) words, the colon here is used to introduce the lists of the stance noun *conclusions*, which, in turn, take the form of the coordinated *that*-clauses. In contrast, comma tends to introduce its following lexico-grammatical elements that explain or amplify the information prior to it, but it does so usually in a way that connects two nominal structures to mark the latter as the apposition of the former (e.g., *the root, the only efficient part*). In other words, it signals the identity meaning relationship between its preceding nominal structure and following one. The same is true for its use as post-modification, as illustrated in Example (11):

(11) Based on charge symmetry**,** [that the masses of the *u* and *d* quarks are much lighter than any other scale in the proton], interchanging $u \leftrightarrow d$ and $\bar{u} \leftrightarrow \bar{d}$ distributions gives the structure functions for scattering from a neutron.

[NSS INT RA26]

In this case, the proposition in the *that*-complement clause "the masses of the *u* and *d* quarks are much lighter than any other scale in the proton", which is introduced by a comma, clarifies the identity or provides descriptive information that helps to specify the stance noun *symmetry*.

The use of dashes is similar to that of commas, but less formal in academic written discourse (Hutauruk, 2016; Sanchez-Stockhammer, 2016), or as Seely (2007: p. 84) puts it, dashes are "the least formal" way of showing apposition. Example (12) is typical in this pattern of use:

DOI:10.6814/NCCU202100169

(12) Via field surveys of attendees at a multiday professional sporting event (n = 1,089), the authors contribute an interesting finding — [that the anticipation of participating in an event sponsor's exhibit area is just as great as the experience itself when it comes to evaluating the sponsor].

[SSS INT RA15]

Finally, the lexico-grammatical features of adverb phrases and extraposition also demonstrate statistically significant differences in their uses. The former functions either to refer head nouns to other parts of the text (i.e., *endophoric marker*) (*RR* = 16.07, very large relative effect size/variance), or to introduce the addition information on stance nouns through the explanation, elaboration, or rephrasing in complement propositions (i.e., *code gloss*) (*RR* = 1.19, small relative effect size/variance). As can be seen in the following example, the adverb phrase *namely* signals or gives readers the cue to the additional information concerning the *theory*, i.e., "that information is used as a mechanism for explaining 'bad news'". The use of *namely*, according to Hyland (2005b), is to guarantee readers' capability of "recover[ing] the writer's intended meaning" (p. 52). In other words, it is used based on the writer's prediction about readers' familiarity to the specific field in question.

(13) On the other hand Wagenhofer (1990) unsuccessfully tested the opposite hypothesis based on signalling theory **namely** [that information is used as a mechanism for explaining 'bad news'].

[NSS INT RA9]

In this regard, it is worth noting that one occurrence of such use in the local subcorpus is particularly distinctive in its unusual position of the complement in the sentence, as illustrated in Example (14):

(14) In recent years, visible light communication (VLC) has drawn increasing attention due to several advantages **such as** high security, low cost, immunity to electromagnetic interference, and [that it is license-free] [4–6].

[NSS LCL RA27]

DOI:10.6814/NCCU202100169

The stance noun *advantages*, in this example, is far distanced from its complement "that it is license-free". In fact, this *that*-clause specifies one of the advantages of visible light communication (VLC), namely, being license-free. This advantage, along with three other advantages which take the form of noun phrases (i.e., high security, low cost, and immunity to electromagnetic interference), constitutes the additional information provided to elaborate the preceding stance noun *advantages*. Such a sentence structure, where a *that*-complement clause is used to provide part of the extra information introduced by *such as* to specify the stance noun, has never been identified in any other research articles of either international or local journals in the corpus data. Typically, the complement immediately follows the code gloss, as demonstrated in the example below also from the local subcorpus:

(15) Note that different occurrences can have overlaps, **namely**, [that two occurrences of a subtree T (T_1 and T_2) may share vertices].

[NSS LCL RA24]

As to the use of endophoric markers, Hyland (1998a: p. 443) defines these metadiscourse devices as "expressions that refer to other parts of the text" (e.g., *as noted above*, *see below*). Endophoric is, therefore, a superordinate term for *anaphoric* (referring backward, i.e., to certain preceding text) and *cataphoric* (referring forward, i.e., to certain subsequent text). Example (16) is typically found in the corpus data:

(16) This trend, which is illustrated in Figure 8, is consistent with the conclusion **above** [that the transition states for decarboxylation and D-exchange reactions show similar stabilizing interactions with the dianion gripper side chains of ScOMPDC over the shared portions of these reactants].

[NSS INT RA8]

Here, the endophoric marker *above* is used to anaphorically refer the stance noun *conclusion* to the previous text.

In the latter case, the use of extraposition reflects writers' adherence to the information principle and the principle of end-weight (*RR* = 3.94, large relative effect size/variance). Extraposition is a lexico-grammatical feature whose syntactic construction involves the

DOI:10.6814/NCCU202100169

alteration of word order in the way that the long and complex grammatical elements are placed towards the end of a clause (see, e.g., Culicover & Rochemont, 1990; Francis, 2010; Wittenberg, 1987), as illustrated in the following example:

(17) Warnings **have also been raised** [that the world is approaching the limit in terms of water availability for agriculture].

[NSS INT RA18]

In this case, the long and complex grammatical element is the *that*-complement clause "that the world is approaching the limit in terms of water availability for agriculture" of the stance noun *warnings*, which has, thus, been placed to the end of the whole sentence. In this regard, the stance noun and its complement are separated by the rest of the sentence constituents, which, though disruptive to the continuity within the *Noun Complement* construction, can be justified by the general principles for the word order within the clause: namely, the information principle and the principle of end-weight (see, e.g., Hasselgård et al., 1998; Tomlin, 2014). According to the information principle, the distribution of information is preferred to be a gradual increase in information load (the amount of new information; cf. given or background information), while the preferred distribution, under the principle of end-weight, corresponds to the weight of grammatical elements. That is, long and complex elements tend to be placed at the end of a clause. Given "heavy" grammatical elements normally bearing a considerable new information load, these two principles often reinforce each other.

4.3 International vs. local writers in the social sciences

As can be seen in Figure 4.2 below, similarities exist between the use of post-modification by social scientists in the international and local journals: (1) that the use of post-modification is most frequently presented by the lexico-grammatical feature *prepositional phrase* in both international and local subcorpora (5.74 vs. 3.71), followed by *punctuation* and *adverb phrase* (same frequency of *adverb phrase* and *noun phrase* in the local subcorpus, both 0.44), and (2) that writers in the social sciences make the least use of *adjective phrase* as post-modification (0.13 vs. 0.04). Differently, *noun phrase* is

DOI:10.6814/NCCU202100169

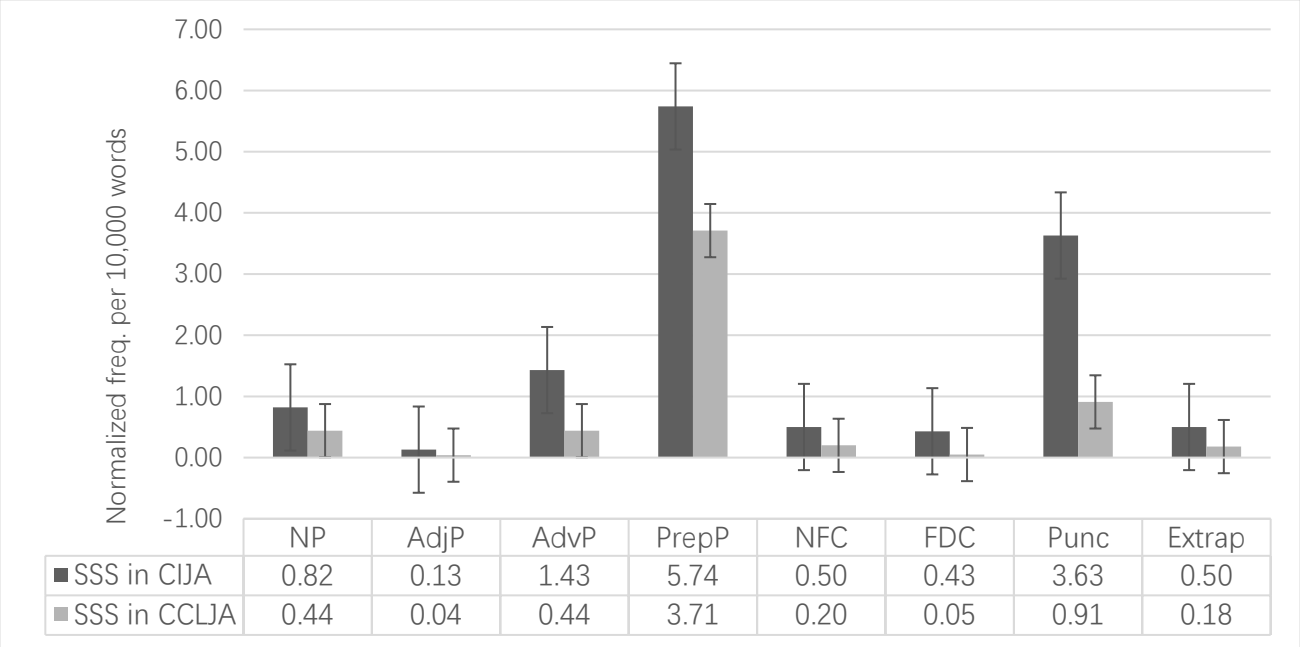


Figure 4.2 Lexico-grammatical features of post-modification use between the Social Scientific Subcorpora (SSSs) in CIJA and CCLJA (normalized frequency per 10,000 words)

Feature	Mann-Whitney			Effect size (*r*)
	U	Z	*p*	
Noun Phrase	**8082.5**	**-4.843**	**< .001**	***r* = -0.280**
Adjective Phrase	**9279.5**	**-4.705**	**< .001**	***r* = -0.272**
Adverb Phrase	**5947.5**	**-7.535**	**< .001**	***r* = -0.435**
Prepositional Phrase	**7066.0**	**-5.572**	**< .001**	***r* = -0.222**
Non-finite clause	**7486.0**	**-6.031**	**< .001**	***r* = -0.348**
Finite dependent clause	**7412.0**	**-7.048**	**< .001**	***r* = -0.407**
Punctuation	**3331.5**	**-10.681**	**< .001**	***r* = -0.617**
Extraposition	**7853.0**	**-5.558**	**< .001**	***r* = -0.221**
Totals	**4158.0**	**-9.441**	**< .001**	***r* = -0.545**

Table 4.3 Lexico-grammatical features of post-modification in the *Noun Complement* construction between SSSs in CIJA and CCLJA

DOI:10.6814/NCCU202100169

the fourth favored lexico-grammatical feature by international writers in their post-modification use, and then their preferences become gradually slighter as to the use of *non-finite clause*, *extraposition* (same frequency as *non-finite clause*, both 0.50), and *finite dependent clause*. On the other hand, local writers choose *non-finite clause* as their fourth preferred lexico-grammatical feature for post-modification use, while also displaying increasingly greater reluctance to access *extraposition* and *finite dependent clause* in their post-modification use. It should be noted here that the international subcorpus contains a significantly higher frequency of post-modification use compared to that in the local subcorpus for each type of lexico-grammatical features.

Inferential statistically, although a significant inferential statistical difference with a medium effect size is reported in terms of the total frequency of post-modification use (Mann-Whitney: U = 4158.0, Z = -9.441, $p < .001$, Effect size: $r = -0.545$), such differences between these two groups of writers with the Pearson *r* value of at least ±.300 (medium effect size) were just found in their use of *punctuation* ($r = -0.617$, $p < .001$), *adverb phrase* ($r = -0.435$, $p < .001$), *finite dependent clause* ($r = -0.407$, $p < .001$), and *non-finite clause* ($r = -0.348$, $p < .001$) (see Table 4.3), which is, in turn, the focus of the analyses on the differences between these two groups of writers.

Moreover, international and local social scientists put their lexico-grammatical features of post-modification into a wider range of functional uses than might be supposed by their natural scientific counterparts (see Table 4.4 below). Specifically, the most statistically significant difference lies in their use of punctuations as post-modification. In this regard, the use of colons is extended to include linking the head noun and its dependent *wh*-interrogative clause (*RR* = 1.28, small relative effect size/variance), as well as the head noun with the substantive meaning and its complement (*RR* = 1.98, medium relative effect size/variance), along with the function of signaling the imminent complex grammatical structures previously identified in the natural scientific subcorpora (*RR* = 1.94, medium relative effect size/variance). Example (18) demonstrates how a colon is used as post-

DOI:10.6814/NCCU202100169

modification to introduce the finite dependent clause which is, regularly, marked by a clause link, namely, a *wh*-word:

(18) For instead of engaging Axis One and Axis Two, they pose only a single question**:** [whether those axes should be sidelined altogether, such that the search or seizure is either reasonable or unreasonable per se, regardless of its factual underpinnings]. [SSS INT RA37]

SSSs in CIJA & CCLJA		
Feature	**Function**	**Variance**
Punctuation		
Colon	Complex grammatical structures	−
	Dependent *wh*-clause	+++++
	Substantive expression	+++++
Comma	Identity relationship	−
	Strategic effort	+++++*11 +
Dash	Informality	-
Semi-colon	Elaborative interpretation	++
Adverb phrase	Interactive metafunctions	−
	Transition marker	− −
	Frame marker	−
	Endophoric marker	− − − − −*3 -
	Code gloss	−
	Interpersonal metafunctions	+++++*2 +
	Hedge	++
	Booster	+++++*3 ++++
	Attitude marker	+++++*3 +++
Finite dependent clause	Abstract entity	+++++*14 +
	Other human	− − − − −*1 −
	Self-sourced	+++++*9 +
	Self-sourced + endophoric marker	+++++*10 +
Non-finite clause	Endophoric marker	+++++*2 ++++
	Abstract entity	− − −
	Other human	−

Note: Relative risk ratio of SSS in CIJA to SSS in CCLJA (minus = SSS in CCLJA to SSS in CIJA).

Table 4.4 Lexico-grammatical features, functions and variance of post-modification use between SSSs in CIJA and CCLJA

DOI:10.6814/NCCU202100169

Also, as can be seen in Example (19), the colon is used to connect the head noun and its complement, given the substantiveness of the head noun, or in other words, the expression in head noun position essential to the delivery of core messages in the research article to its readers. Specifically, the head noun *principle*, as the author put it, is "deeper" and contributes to the understanding of the "first four principles", thereby rendering it essential to the authors' complete explanation of the other four principles.

(19) The first four principles are all motivated by one deeper principle**:** [that sin taxes should be designed to offset uninternalized harms].

[SSS INT RA46]

The co-text of Example (19) below can provide a clearer picture of how the substantive expression *principle* relates to the authors' explanatory process:

Guiding Principles for Policymakers

Although uncertainty remains about some empirical parameters, economic theory and existing data suggest seven guiding principles for designing sugar-sweetened beverage taxes. The first four principles are all motivated by one deeper principle: that sin taxes should be designed to offset uninternalized harms.

1. Focus on Counteracting Externalities and Internalities, Not on Minimizing Sugary Drink Consumption

Many public health advocates explicitly or implicitly take the perspective that the goal of policymakers should be to maximize health or minimize unhealthy behaviors. It's easy to see why this can't be the right social objective. The way to maximize health is to ban any sugary or fatty food or drink, including sugary drinks, red meat, and dessert. Such a ban would preclude any enjoyment that people get from eating steak or dessert, and it's not clear where to draw the line on what foods or drinks to ban.

… (bold in the original)

In this excerpt, immediately after conveying what the "deeper" *principle* is (Example (19) underlined), the article unfolds with the detailed explanations for each of the first four

DOI:10.6814/NCCU202100169

principles in sequence based on this "deeper" *principle*, which is, in turn, rendered substantive in the expression of meaning in the text.

Additionally, commas are used for writers' launch of strategic effort (*RR* = 1.94, medium relative effect size/variance), and also serve to signal an identity meaning relationship (*RR* = 3.50, large relative effect size/variance). In terms of the newly identified function of launching writers' strategic effort, to illustrate, the head noun *factors*, in Example (20), is separated from its coordinated *that*-complement clauses by a comma, which is, however, not used to simply indicate that its following complement is about to provide descriptive information for the *factors*:

(20) But the differing results could be also be due to a number of other factors**,** [including that the survey data has fewer observations and that the Kuhn and Weidmann (2015) analysis uses a different set of control variables].

[SSS INT RA11]

Although in complement position the writer merely list two other factors in the form of coordinated *that*-clauses, it is reasonable to assume that their number is more than two, since the above excerpt states "*a number of* other factors". In this sense, two scenarios are thus proposed here: first, that the writer just exemplifies two most significant other factors while ignoring the others, and second, that the writer skillfully chooses the two other factors that contribute to the justification of his or her point of view. Provided the second scenario is confirmed, the writer here is making strategic effort through the use of a comma, which, in turn, separates the head noun and its complement, thus creating a certain leeway for writers to choose the meaning elements to their benefit (e.g., only choosing the factors beneficial to their argumentation as in the above example). In so doing, the probability of success in bringing readers into alignment with their stances is likely to increase.

The function of dashes remains unchanged (*RR* = 1.37, small relative effect size/variance), while a new feature category of *semi-colon* is added for elaborative interpretations (*RR* = 1.30, small relative effect size/variance), as in this example:

DOI:10.6814/NCCU202100169

(21) When trial judges are left with broad discretion to resolve unfamiliar questions of law in complex circumstances, there is a risk they will latch onto rubrics that are a poor fit**;** [that individual judges' inability to evaluate foreign facts and systemic interests as part of routine procedural decisions will increasingly ossify, generalize, and ultimately marginalize factors initially meant to protect comity].

[SSS INT RA29]

Here, as "the semi-colon allows elaborative interpretations" (Huddleston & Pullum, 2002: p. 1742), the *that*-complement clause "that individual judges' inability to evaluate … will increasingly ossify, generalize, and ultimately marginalize factors initially meant to protect comity" immediately following this punctuation mark is used to elaborate or interpret the stance noun *fit* that precedes it.

Moreover, the functions of adverb phrases have been further divided into interactive (i.e., transition markers, frame markers, endophoric markers, and code glosses) and interpersonal metafunctions (i.e., hedges, boosters, and attitude markers) in the social scientific subcorpora. Whereas the former aids readers in comprehending a text by explaining, orienting and guiding them through the information (*RR* = 1.22, small relative effect size/variance), the latter helps writers to express their attitudes and judgements about a text or discourse participants, and characterize the interaction between writer and reader (the writer-reader relationship) (*RR* = 4.41, large relative effect size/variance). To illustrate, the new function of adverb phrases as transition markers found in the social scientific subcorpora is to show relations between main clauses, as illustrated in the following example:

(22) Discussions of nonbinary gender rights are often stifled by the assumption that those rights must always take the form of gender neutrality or, **alternatively**, [that the law must always recognize a third gender].

[NSS INT RA41]

In this extract, the transition marker *alternatively* functions to indicate the relation between the two *that*-complement clauses as comparison. Another new function of adverb

DOI:10.6814/NCCU202100169

phrases pertains to its linguistic role as frame markers, which are typically used for the signal of sequencing (e.g., *firstly*), the label of text stages (e.g., *overall*, *so far*), the announcement of discourse goals (e.g., *(in) this part, (in) this chapter*), or the indication of topic shifts (e.g., *now*, *right*), for example:

(23) The declaration sets out two key principles **firstly** [that authors grant "to all users a free, irrevocable, worldwide, right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship"].

[SSS INT RA28]

The frame marker *firstly*, in this example, is placed between the head noun *principles* and its *that*-complement clause, in a way that indicates the following complement being the first one in the sequences vis-à-vis all other principles included in the subsequent text.

As to the interpersonal metafunctions, hedges (e.g., *mostly*, *perhaps* and *possibly*) are the devices used by writers to show their withdrawal of full commitment to the truth value of the proposition, also indicating their acknowledgement of alternative voices and positions (Ädel, 2006; Biber, 1991; Hyland, 2005b, 2014). As in the following example, the hedge *possibly* serves the function of withholding a complete commitment to the truth of the proposition nominalized into the head noun *reason* and specified in its *that*-complement clause "that the grandest and simplest of his ideas seemed to him such an absolute matter of plain, common-sense". In other words, *possibly* implies that the *reason* is based on the writer's plausible reasoning rather than accredited facts, showing the writer's degree of confidence in presenting it. However, from another perspective, it also demonstrates the writer to be open-mined about others' proposals for alternative reasons.

(24) Indeed, none of them had ever known anything else, as Napoleon had never explained to him the real mechanism of his thoughts, for the reason **possibly** [that the grandest and simplest of his ideas seemed to him such an absolute matter of plain, common-sense].

DOI:10.6814/NCCU202100169

[SSS INT RA28]

In contrast, boosters (e.g., *certainly*, *evidently* and *absolutely*) allow writers to convey certainty in what they say, thereby increasing the force of the proposition and displaying a strong commitment to it (Hyland, 2004, 2005b, 2009; Martin & White, 2005). In the typical case below, the writer employs the booster *certainly* to underline the conviction he wishes to attach to the *argument*:

(25) There is an argument **certainly** [that one of the key driving forces behind the Chinese culture change is the ecological and environmental divisions between southern and northern China].

[SSS LCL RA63]

Attitude marker, when taking the form of adverb phrases (e.g., *interestingly*, *surprisingly* and *unexpectedly*), is another interpersonal metadiscourse device placed in the position between a head noun and its complement. It usually functions to signal the writer's affective, rather than epistemic, attitudes to the propositional content, conveying surprise, agreement, importance, frustration, etc., but not commitment (Hyland, 2004, 2005b, 2009; Martin & White, 2005), as in Example (26):

(26) The following finding with nasalised vowel î overruled my assumption **surprisingly** [that high vowels will only take another high vowel to form a permissible combination].

[SSS INT RA74]

Here, the use of the adverb *surprisingly* as the attitude marker signals the writer's affective attitude of surprise to the head noun *assumption* with its specification in the following *that*-complement "that high vowels will only take another high vowel to form a permissible combination". Put differently, the attitude marker *surprisingly* is deployed here to convey the writer's surprise on the rejection of his previous assumption.

In addition, finite dependent clauses are devoted to indicating whether the writer explicitly avers the stance noun (see Example (27)) (*RR* = 18.30, very large relative effect size/variance), or attributes it to other humans (see Example (28)) (*RR* = 3.15, large relative

DOI:10.6814/NCCU202100169

effect size/variance) or abstract entities (see Example (29)) (*RR* = 28.80, very large relative effect size/variance).

(27) These observations lend support to the theory **which we are advancing** [that ionization, clue directly to ultraviolet light or indirectly to the ionized condition of the gas (air) which forms the medium surrounding the leaf, is responsible for the potentiometric changes observed].

[NSS INT RA43]

(28) First, my claim that the mind is a computational system is different from the claim **Fodor attacks** [that the mind has the architecture of a Turing Machine].

[NSS INT RA35]

(29) The underlying assumption **on which these models are based** [that the different groups of rice could not have acquired domestication alleles from standing variation in the wild population] is clearly incorrect.

[NSS LCL RA18]

They also serve the newly identified function of conveying the writer's averral combined with the expression referring to other parts of the text (*self-sourced + endophoric marker*; see Example (30)) (*RR* = 21.20, very large relative effect size/variance).

(30) A consequence of this autonomy in the decision-making structure is the fact **which we noted above** [that an ASO is responsible primarily to the HMCs in his area].

[SSS INT RA16]

Here, the relative-clause subject and verb *we noted* is drawn on by the writers to indicate their averral of the stance expressed through the head noun *fact* and specified in its *that*-complement clause "that an ASO is responsible primarily to the HMCs in his area" (i.e., *self-sourced* stance-making), while the adverbial position within this clause is occupied by an adverb *above*, which, then, functions as the endophoric marker to anaphorically refer the *fact* back to the previous text.

DOI:10.6814/NCCU202100169

In the parallel vein, non-finite clauses serve the functions of *other human* (see Example (31)) (*RR* = 1.26, small relative effect size/variance), *abstract entity* (see Example (32)) (*RR* = 1.69, small relative effect size/variance) and *endophoric marker* (see Example (33)) (*RR* = 5.88, large relative effect size/variance).

(31) In spite of cautiously proposed rules of thumb available in the PLS literature, we are frustrated by sweeping claims **made by some researchers** [that PLS modeling can or should be used (often instead of the covariance-based approach) because it makes no sample size assumptions or because somehow "Sample size is less important in the overall model" (Falk and Miller 1992, p. 93)].

[NSS INT RA50]

(32) In this case, the national authorities were free to provide for an appeal with suspensive effect against decisions **relating to an EAW** [to uphold domestic constitutional rights], because the EAW Framework Decision left them a degree of discretion in this regard.[105]

[NSS INT RA81]

(33) Then we shall use this theory **to show in Section III** [that the one-period version of

(1) $$V = V_u + \tau B$$

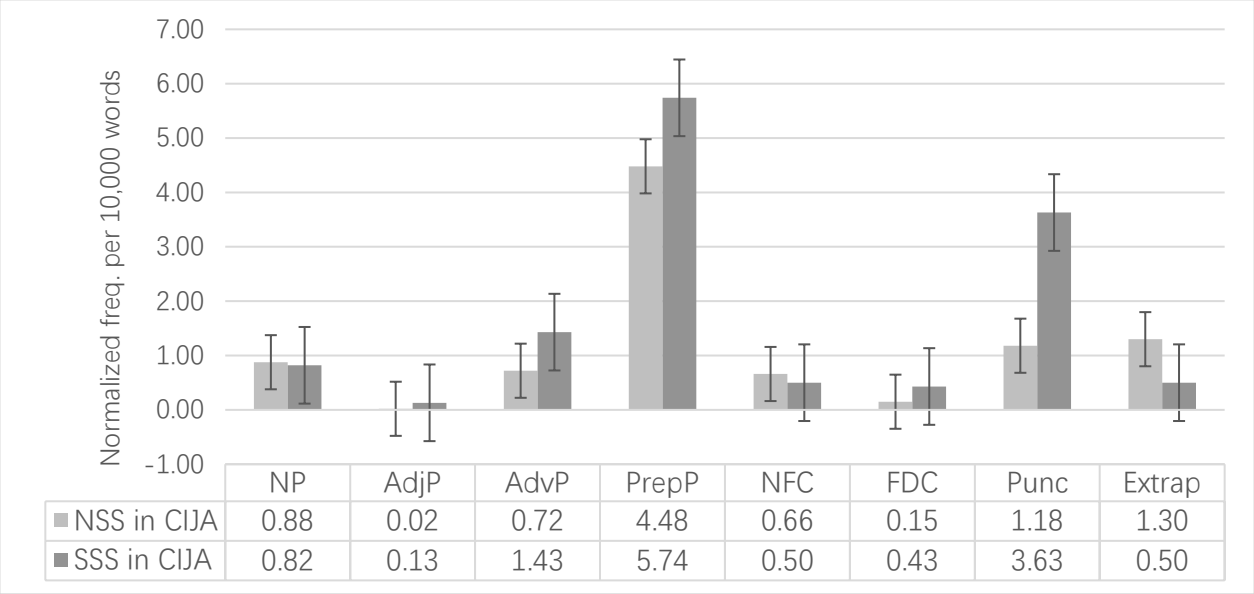
is a "short-run" equation1 that would bear no resemblance to the equation faced by the firm and its competitors in the "long run."]

[NSS INT RA23]

4.4 International writers in the natural vs. social sciences

International natural and social scientists' post-modification use differs greatly in the frequencies of its lexico-grammatical features (see Figure 4.3 below). Generally, both natural and social subcorpora contain the highest frequency of *prepositional phrase* as post-modification, with that of *adjective phase* being the lowest. More specifically, writers from the international journals of the natural sciences exhibit their second preference for *extraposition*, followed by *punctuation*, *noun phrase*, *adverb phrase*, *non-finite clause* and

DOI:10.6814/NCCU202100169



	NP	AdjP	AdvP	PrepP	NFC	FDC	Punc	Extrap
NSS in CIJA	0.88	0.02	0.72	4.48	0.66	0.15	1.18	1.30
SSS in CIJA	0.82	0.13	1.43	5.74	0.50	0.43	3.63	0.50

Figure 4.3 Lexico-grammatical features of post-modification use between NSS and SSS in CIJA (normalized frequency per 10,000 words)

Feature	Mann-Whitney			Effect size (*r*)
	U	Z	*p*	
Noun Phrase	**8848.5**	**-3.627**	**< .001**	***r* = -0.209**
Adjective Phrase	**9015.5**	**-5.551**	**< .001**	***r* = -0.121**
Adverb Phrase	**6817.5**	**-6.299**	**< .001**	***r* = -0.357**
Prepositional Phrase	**8246.0**	**-4.002**	**< .001**	***r* = -0.231**
Non-finite clause	**8941.0**	**-3.559**	**< .001**	***r* = -0.205**
Finite dependent clause	**8093.5**	**-5.581**	**< .001**	***r* = -0.322**
Punctuation	**3879.0**	**-9.971**	**< .001**	***r* = -0.576**
Extraposition	9884.0	-1.988	= .047	***r* =** -0.115
Totals	**7506.0**	**-4.984**	**< .001**	***r* = -0.288**

Table 4.5 Lexico-grammatical features of post-modification in the *Noun Complement* construction between NSS and SSS in CIJA

DOI:10.6814/NCCU202100169

finite dependent clause. By comparison, their counterparts in soft fields have *punctuation* as their second choice, with a gradually vague possibility of choosing *adverb phrase*, *noun phrase*, *non-finite clause*, *extraposition* (same frequency as non-finite clauses, both 0.50), and *finite dependent clause* in their post-modification use. Table 4.5, moreover, demonstrates the inferential statistical differences in how international writers in the soft and hard sciences deploy such lexico-grammatical features in their post-modification use. In this regard, although the difference in the total frequency of post-modification use (Mann-Whitney: U = 7506.0, Z = -4.984, $p < .001$, Effect size: $r = -0.288$) is not that notable given its nearly medium effect size ($|r| > .300$) found, the employment of *punctuation*, in light of the large effect size ($|r| > .500$) reported ($r = -0.576$, $p < .001$), demonstrates a particularly noticeable statistical difference between international natural and social sciences writers, followed by the medium effect sizes obtained for *adverb phrase* ($r = -0.357$, $p < .001$) and *finite dependent clause* ($r = -0.322$, $p < .001$), respectively. As such, this research is going to focus on the analyses of these three lexico-grammatical features of international natural and social scientists' post-modification use, since only linguistic features with a Pearson *r* value of at least ±.300 (medium effect size) are, as noted above, qualified for inclusion of the current study.

International hard and soft field writers' discipline-specific preferences are also reflected in their use of the lexico-grammatical features with significant statistical differences to perform a variety of intended functions (see Table 4.6 below). Specifically, colons in the feature category of punctuation function to introduce complex grammatical structures (*RR* = 1.58, small relative effect size/variance), dependent *wh*-interrogative clauses (*RR* = 12.30, very large relative effect size/variance) and the elaboration of substantive expressions (*RR* = 24.30, very large relative effect size/variance), while commas are for identity relationships (*RR* = 1.29, small relative effect size/variance) and strategic effort (*RR* = 22.20, very large relative effect size/variance), with dashes marking informality (*RR* = 2.02, small relative effect size/variance) and semi-colons signalling

DOI:10.6814/NCCU202100169

NNS & SSS in CIJA		
Feature	**Function**	**Variance**
Punctuation		
Colon	Complex grammatical structures	− − −
	Dependent *wh*-clause	+ + + + +*6 +
	Substantive expression	+ + + + +*12 +
Comma	Identity relationship	−
	Strategic effort	+ + + + +*11 +
Dash	Informality	+ + + + +
Semi-colon	Elaborative interpretation	+ +
Adverb phrase	Interactive metafunctions	−
	Transition marker	− − − − −*7 -
	Frame marker	− − − − −*4 -
	Endophoric marker	+ + + + +*16 +
	Code gloss	+ + +
	Interpersonal metafunctions	+ + + + +*11 +
	Hedge	+ + + + +*3 +
	Booster	+ + + + +*3 +
	Attitude marker	+ + + + +*3 +
Finite dependent clause	Abstract entity	− − − − −*1 − −
	Other human	+ + + + +*15 + + + +
	Self-sourced	+ + + + +*9 +
	Self-sourced + endophoric marker	+ + + + +*10 +

Note: Relative risk ratio of SSS in CIJA to NSS in CIJA (minus = NSS in CIJA to NSS in CIJA).

Table 4.6 Lexico-grammatical features, functions and variance of post-modification use between NNS and SSS in CIJA

elaborative interpretations (*RR* = 1.30, small relative effect size/variance). Their deployment of adverb phrases as post-modification to carry out interactive (*RR* = 1.29, small relative effect size/variance) and interpersonal metafunctions (*RR* = 22.50, very large relative effect size/variance) also demonstrate statistically significant disciplinary differences. Finally, the functions of *abstract entity* (*RR* = 3.47, medium relative effect size/variance), *other human* (*RR* = 31.70, very large relative effect size/variance), *self-sourced* (*RR* = 18.30, very large relative effect size/variance) and *self-sourced* +

DOI:10.6814/NCCU202100169

endophoric marker (*RR* = 21.20, very large relative effect size/variance) are fulfilled through the employment of finite dependent clauses.

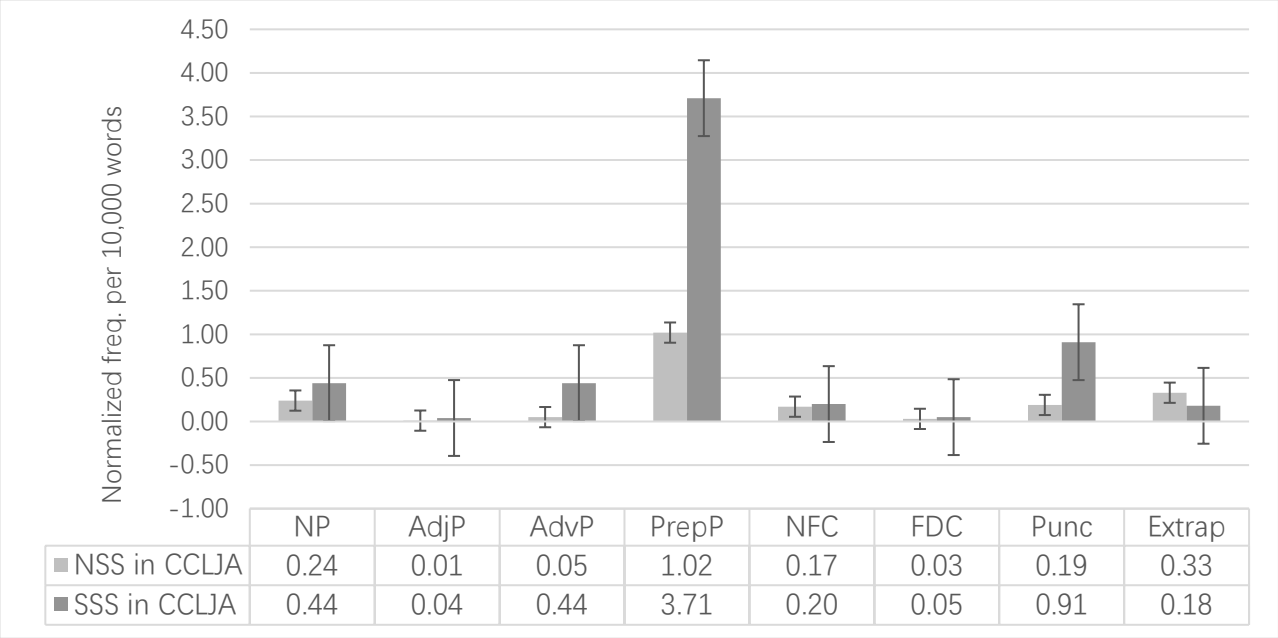
4.5 Local writers in the natural vs. social sciences

There are quite different frequencies of grammatical elements used as post-modification in the *Noun Complement* construction by writers from local journals in the hard and soft fields (see Figure 4.4 below). When it comes to the points of similarity, both local natural and social scientists most prefer *prepositional phase* in their post-modification use, whereas *adjective phrase* turns out to be their least popular choices, preceded by *finite dependent clause* in both subcorpora. As to the patterns of difference, *extraposition* becomes the second preference of writers in hard science journals, followed by *noun phrase*, *punctuation*, *non-finite clause*, and *adverb phrase*, while local soft scientists draw on *punctuation* in the second highest frequency, with their enthusiasm tempering from *noun phrase*, *adverb phrase* (same frequency of adverb phrases and noun phrases in local subcorpus, both 0.44), *non-finite clause* to *extraposition* as post-modification.

More importantly, significant inferential statistical differences were revealed between local writers from soft and hard fields in their preferences for certain grammatical forms in post-modification use (see Table 4.7 below). Although in general, the difference pertaining to the total frequency of post-modification use (Mann-Whitney: U = 4581.5, Z = -9.062, $p < .001$, Effect size: $r = -0.523$), given the large effect size reported ($|r| > .500$), is clearly discernible, the statistical differences are particularly noticeable merely for two lexico-grammatical features in the two subcorpora – *prepositional phrase* with the large effect size found ($r = -0.502$, $p < .001$) and medium effect sizes reported for *punctuation* ($r = -0.380$, $p < .001$). As such, the remainder of this study will focus on just these two grammatical forms as post-modification in the *Noun Complement* construction used by local natural and social scientists.

Functionally, prepositional phrases, as illustrated in Table 4.8 below, are used by these two groups of published expert writers on a disciplinarily different basis to serve the functions of *self-sourced* (*RR* = 1.10, small relative effect size/variance), *other human* (*RR*

DOI:10.6814/NCCU202100169



	NP	AdjP	AdvP	PrepP	NFC	FDC	Punc	Extrap
NSS in CCLJA	0.24	0.01	0.05	1.02	0.17	0.03	0.19	0.33
SSS in CCLJA	0.44	0.04	0.44	3.71	0.20	0.05	0.91	0.18

Figure 4.4 Lexico-grammatical features of post-modification use between NSS and SSS in CCLJA (normalized frequency per 10,000 words)

Feature	Mann-Whitney			Effect size (*r*)
	U	Z	*p*	
Noun Phrase	**9765.5**	**-3.243**	**= .001**	***r* = -0.187**
Adjective Phrase	11026.0	-1.345	= .179	***r* =** -0.078
Adverb Phrase	**10168.0**	**-2.176**	**= .012**	***r* = -0.157**
Prepositional Phrase	**4982.5**	**-8.696**	**< .001**	***r* = -0.502**
Non-finite clause	**10328.5**	**-2.392**	**= .017**	***r* = -0.138**
Finite dependent clause	10803.5	-2.130	= .033	***r* =** -0.123
Punctuation	**7152.0**	**-7.009**	**< .001**	***r* = -0.405**
Extraposition	11099.5	-0.347	= .729	***r* =** -0.020
Totals	**4581.5**	**-9.062**	**< .001**	***r* = -0.523**

Table 4.7 Lexico-grammatical features of post-modification in the *Noun Complement* construction between NSS and SSS in CCLJA

DOI:10.6814/NCCU202100169

NNS & SSS in CCLJA		
Feature	**Function**	**Variance**
Prepositional phrase	Self-sourced	+
	Other human	+++++*11 +++
	Abstract entity	----
	Qualifier + subject	++
Punctuation		
Colon	Complex grammatical structures	+++++*40 ++
	Dependent *wh*-clauses	+++++*3 ⊣
	Substantive expression	+++++*6 ⊣
Comma	Identity relationship	-
Dash	Informality	--

Note: Relative risk ratio of SSS in CCLJA to NSS in CCLJA (minus = NSS in CCLJA to SSS in CCLJA).

Table 4.8 Lexico-grammatical features, functions and variance of post-modification use between NNS and SSS in CCLJA

= 23.67, functions of *self-sourced* (*RR* = 1.10, small relative effect size/variance), *other human* (*RR* = 23.67, very large relative effect size/variance), *abstract entity* (*RR* = 1.88, small relative effect size/variance), and *qualifier + subject entity* (*RR* = 1.32, small relative effect size/variance). The feature category of punctuation assigns colons to deal with complex grammatical structures (*RR* = 81.30, very large relative effect size/variance), dependent *wh*-interrogative clauses (*RR* = 6.20, large relative effect size/variance) and substantive expressions (*RR* = 12.50, very large relative effect size/variance), commas to signal identity relationships (*RR* = 1.08, small relative effect size/variance), and dashes to convey informality (*RR* = 1.36, small relative effect size/variance).

DOI:10.6814/NCCU202100169

CHAPTER 5

Discussion

5.1 Introduction

The previous chapter reports the different use of the lexical-grammatical features and functions of the target structure between disciplinary fields and writers of contrasting linguistic-cultural backgrounds. Based on these findings, the goals of this chapter are thus to discuss the underlying motivations behind these variations, demonstrate the new perspectives and revelations concerning the stance-making roles the post-modification structures play in academic writing, and make comparisons between this dissertation study and prior research to further discuss why the use of post-modification structures is worth looking at more closely in this line of research. The explanatory model of cross-disciplinary and cross-cultural differences presented in the methodology chapter was adopted to conduct the reason analyses. The discussion of the underlying motivations are presented in order of the comparisons between (a) international and local writers in the natural sciences, (b) international and local writers in the social sciences, (c) international writers in the natural and social sciences, and (d) local writers in the natural and social sciences.

5.2 Explanations for different use between international and local writers in the natural sciences

There are significant statistical differences (Pearson's $r \geqq \pm.300$) in the use of prepositional phrases, punctuation, adverb phrases and extraposition as post-modification

Feature	Mann-Whitney			Effect size (r)
	U	Z	p	
Prepositional Phrase	**5107.0**	**-8.558**	**< .001**	**r = -0.494**
Punctuation	**7506.5**	**-6.586**	**< .001**	**r = -0.380**
Adverb Phrase	**8328.0**	**-6.261**	**< .001**	**r = -0.362**
Extraposition	**7631.0**	**-6.110**	**< .001**	**r = -0.353**

Table 5.1 Lexico-grammatical features of significant inferential statistical differences (in descending order of effect size *r* value) between NSSs in CIJA and CCLJA

DOI:10.6814/NCCU202100169

	NSS in CIJA	NSS in CCLJA
Self-sourced	3 (0.9)	0 (0.0)
Other human	15 (4.5)	1 (1.5)
Abstract entity	242 (72.7)	54 (84.4)
Qualifier + subject	73 (21.9)	9 (14.1)
Totals	333 (100.0)	64 (100.0)

Table 5.2 Prepositional phrase as post-modification between NSSs in CIJA and CCLJA (% of total)

(in descending order of effect size *r* value) between international and local natural scientists (see Table 5.1). Therefore, the analyses of the underlying motivations for the different use of the lexico-grammatical features and functions of the target structure will start with that of prepositional phrases as post-modification, since it demonstrates the most different preferences between these two groups of writers (highest effect size *r* value).

5.2.1 Prepositional phrase

Table 5.2 illustrates the functional characteristics of the prepositional phrase use between the two contrasting subcorpora. Both international and local writers in the natural sciences show the strongest inclination to choose prepositional phrases for the attribution of stance nouns to abstract entities (72.7 vs. 84.4). Their most frequent use of abstract entities can be attributed to the mode of knowledge construction in the natural sciences. That is, knowledge in these fields is built mainly on empirical evidence and origination of facts by virtue of experimentation and scientific observation (Becher & Trowler, 2001; Hockin et al., 1996; Jiang & Hyland, 2015; Lee & Tan, 2018), which, in turn, calls for the attribution of varying evidence or facts to their sources, namely, the abstract entities of different kinds involved in experiments and observations (e.g., the results or research methods), as shown in the following example:

(34) Due to the overlapping signals, it is not possible to definitively distinguish the formation time of this signal; however, it is certainly present by 2-3 ps in CNINDT, effectively ruling out the formation of these triplets from ISC, and confirming the result **from EPR** [that triplets are originating from singlet fission].

DOI:10.6814/NCCU202100169

[NSS INT RA12]

Here, the stance noun *result* refers to certain facts or empirical evidence obtained from the abstract entity *EPR*, and actually there exist many more stance nouns like *result* that need to be attributed to their corresponding abstract entities in the natural science fields. Therefore, the use of prepositional phrases as post-modification for the attribution to abstract entities leaves more room for natural scientists to mark the sources of stance nouns.

In contrast, both international and local natural scientists are most reluctant to overtly aver a stance noun: 3 occurrences (0.9) and 0 occurrences (0.0) in the international and local writers' subcorpora, respectively. This is not surprising since the entrenched conventions of academic writing in English demand impersonality in research articles, which recommends writers to conceal themselves in text and to substitute persuasive objectivity for subjective interpretations (Guinda & Hyland, 2012; Hewings, 2012; Hyland, 2009, 2013). Absence of the *self-sourced* use can be seen particularly in the research articles written by natural scientists, not only because these authors are simply "messenger[s] relaying the truth from nature" (Gilbert, 1976: p. 285), but also because they are exhorted to "strengthen the objectivity of their interpretations and subordinate their own voice to that of unmediated nature" (Hyland, 2001: p. 216). It should be noted here that the international and local subcorpora do demonstrate rather slight differences in the self-sourced use, with none in the latter and three occurrences in the former. In this sense, it seems that local writers tend to perform stricter adherence to the conventions of academic discourse in English, thus becoming more qualified messengers relaying the truth from nature. However, seen from another angle, it is reasonable to infer that local writers may draw on other lexico-grammatical features (e.g., first-person possessives as pre-modifiers) to conduct self-mention, as in the cases reported by other studies where the frequencies of self-mention (e.g., *my* and *our*) in Chinese EFL writers' work are significantly higher (even 9 times) than that of English native equivalence (e.g., Chen, 2014; Jiang, 2015).

On the other hand, international and local writers from the field of natural sciences differ considerably in their preferences as to the rest functional characteristics of this

DOI:10.6814/NCCU202100169

pattern of use. In terms of *other human*, the international subcorpus has exactly three times more of its use than the local one (4.5 vs. 1.5). The other human use, as one type of attribution, normally takes the form of a citation in academic writing. Thompson (2012) further categorizes them as *integral* (placed in the sentence with an explicit grammatical role) and *non-integral citation* (placed outside the sentence, such as, in brackets, without a grammatical role). As to prepositional phrases as post-modification for other human use, it can be seen as belonging to the kind of integral citations, given this post-modification structure normally placed within the sentence while also playing an explicit grammatical role, as shown in Example (35):

(35) That was in line with a previous suggestion **by Kohl** [4] [that turgor changes in guard cells may be related to the hydrolysis and synthesis of starch existing within the guard cells].

[NSS INT RA11]

To further reveal the underlying reasons for the differences in this pattern of use between international and local natural scientists, the co-text of Example (35) (in **bold**) was examined, as follows:

Fifty-two years after the prediction of the osmolyte-turgor mechanism, Francis Darwin reported that light-triggered stomatal opening also occurs in CO2-free air, indicating that stomatal movement does not depend on any immediate products of photosynthetic carbon assimilation as potential osmolytes [3]. **That was in line with a previous suggestion by Kohl [4] that turgor changes in guard cells may be related to the hydrolysis and synthesis of starch existing within the guard cells.** Francis Ernst Lloyd [5] set about to examine the starch theory and, to that end, used iodine staining to monitor the starch content of guard cells over the course of the day. He noticed that the starch-granule content of guard cells was higher in the closed stomata of Verbena ciliata and Ocotillo (the exact species is not specified), but lower in open stomata. He also examined, at various times of the day, several other species (e.g., Amaranthus,

DOI:10.6814/NCCU202100169

	NSS RAs in CIJA	NSS RAs in CCLJA
Average text length (number of words)	6,516	3,834
Average pages per RA	19	10

Table 5.3 Average text length and pages per RA between NSSs in CIJA and CCLJA

Jatropha, Nicotiana, Euphorbia), but he failed to find a similar pattern of changes in starch-granule content that could be related to stomatal opening [5].

In this extract, the whole paragraph is part of the literature review of how the starch-sugar theory was developed. Three researchers who contributed to the development of the theory, namely, *Francis Darwin, Kohl* and *Francis Ernst Lloyd*, are identified through integral citations. In the case of *Kohl*, the prepositional phrase as post-modification is used to refer the stance noun *suggestion* to its source *Kohl*. According to Thompson (2001), integral citations focus on the cited authors. However, it is not characteristic of a natural science research article to contain a detailed literature review section whose focus is on cited authors, given its space/word limit (on average 7 pages/7193 words per article in the two subcorpora). Therefore, it seems that articles with a space of several paragraphs to extensively review a concept, theory and the like are likely to include prepositional phrases as post-modification for other human use. Articles of this kind are not many in the international subcorpus, which thus explains why there is only 4.5% of prepositional phrases as post-modification used to indicate the source of a stance noun as other humans, while they are much fewer in the local one, which can be attributed to the higher space/word limit in local journals (see Table 5.3), therefore accounting for their non-use of the other human sources.

Lastly, natural science writers in international journals hinge far more upon prepositional phrases for the *qualifier + subject* use than their local counterparts. This might be mainly due to the grammatical complexity of this particular *Noun Complement* construction. Specifically, the prepositional phrase in this case, as noted above, functions simultaneously as the qualifier of its preceding stance noun and as the subject of its following proposition. Example (36) is typical in the two subcorpora:

DOI:10.6814/NCCU202100169

(36) This component of NASA's broader Solar System exploration mission is driven in part by the 1998 Congressional directive **to NASA** [to protect the Earth and its inhabitants from the threat of asteroid impact].

[NSS INT RA33]

Here, the content of the stance noun *directive* is specified in its *to-infinitive* complement clause "to protect the Earth and its inhabitants from the threat of asteroid impact". The prepositional phrase *to NASA*, as the qualifier of the stance noun *directive*, limits its meaning to a specific directive issued exclusively for NASA, and meanwhile, it also plays the part of subject in the following *to-infinitive* complement, since it is NASA that protects the Earth and its inhabitants from the threat of asteroid impact. In this respect, the prepositional phrase *to NASA* seems to take on a linking role in making the relationship between the stance noun and the following proposition closer, which, in turn, facilitates readers' understanding of this particular *Noun Complement* construction. However, to draw on such a construction also imposes stricter requirements on the writers, given that they first need to acquire a clear understanding of the relationship among these three sentence constituents. This partly explains why local natural scientists are, perhaps due to their lower grammatical and rhetorical awareness of this specific relationship in English, more reluctant to use this construction.

Furthermore, within the international subcorpus, it is also noticeable that over one third of prepositional phrases in the *qualifier + subject* use are complex ones (see Table 5.4 below), as in the following example:

(37) This clearly incurs a demand **for P in photosynthetic and respiratory metabolism** [to support ATP production].

[NSS INT RA4]

In this excerpt, the specific content of the stance noun *demand* is demonstrated in its complement "to support ATP production", and what it differs from the examples above lies in the complex prepositional phrase within the sentence which consists of two prepositional phrases between the stance noun and its *to-infinitive* complement. In fact, the

DOI:10.6814/NCCU202100169

	NSS RAs in CIJA	NSS RAs in CCLJA
Single prepositional phrase	50 (68.5)	9 (100.0)
Complex prepositional phrase	23 (31.5)	0 (0.0)
Totals	73 (100.0)	9 (100.0)

Table 5.4 Types of prepositional phrase used between NSSs in CIJA and CCLJA (% of total)

second prepositional phrase *in photosynthetic and respiratory metabolism* is the postnominal modifier of the first one *for P* within the whole complex prepositional phrase. In this sense, international writers in the nature sciences not only incline to draw on more prepositional phrases in their *qualifier + subject* use, but they are also far more likely to use them in a grammatically complex way, as compared to the local writers whose use of complex prepositional phrases in the *qualifier + subject* construction is none (see Table 5.4). This is unsurprising given the increasing grammatical complexity of this particular construction with the inclusion of complex prepositional phrases, which, due to local writers' relatively lower levels of proficiency in English, seems to prevent them from making more use of it than their international counterparts. Or, in Ellis's (1994: p. 185) words, this likely means that local writers prefer to avoid using this structure rather than make wrong attempts through the so-called "avoidance strategies".

5.2.2 Punctuation

The second most statistically distinct lexico-grammatical feature of the post-modification use between international and local natural scientists is *punctuation mark*, whose three forms have been identified from the two subcorpora, namely, *comma*, *colon* and *dash*.

In early literature, Crismore et al. (1993) note that punctuation marks are able to "

	NSS in CIJA	NSS in CCLJA
Comma (,)	42 (40.0)	6 (50.0)
Colon (:)	49 (46.7)	0 (0.0)
Dash (—)	14 (13.3)	6 (50.0)
Totals	105 (100.0)	12 (100.0)

Table 5.5 Types of punctuation used between NSSs in CIJA and CCLJA (% of total)

DOI:10.6814/NCCU202100169

signal text glosses and clarifications as well as uncertainty, certainty, and attitude" (p. 48). More recently, Hyland (2019) confirmed their findings by pointing out that punctuation mark in its different forms functions to foreground aspects of written discourse or the author's attitude toward them. Following these, the present analysis is going to focus on not only the three punctuation marks' function of introducing their subsequent information to explain or amplify the one that precedes them (or, in Biber and Gray's terms (2016), indicating the explanatory relationship), but also the writers' extra-linguistic effort beyond this pattern of use.

The punctuation mark *colon* is examined first, given the most distinct difference in its deployment between writers from soft and hard domains (see Table 5.5). Example (38) is typical in its use:

(38) This has led to two related hypotheses**:** [that genome duplication may have conferred an 'extinction resistance' to particular lineages of plants, and that polyploid genomes may have allowed surviving lineages to rise to dominance in the wake of this mass extinction episode].

[NSS INT RA23]

In this excerpt, the colon placed between the stance noun *hypotheses* and its *that*-complement clauses indicates that explanations are about to come, which may, to some extent, facilitate readers' understanding of the meaning relation between stance noun and complement clauses (i.e., explanatory relationship in this case). This typically occurs when the author predicts that the readers are highly likely to find it difficult or problematic to figure out the meaning relation between a stance noun and the complement clauses of considerable grammatical complexity (e.g., the coordinated *that*-clauses as in this case). Example (39) is also typical in the international writers' subcorpus:

(39) With these analyses, we tested several specific predictions**:** [(1) that points of muscle insertion should be centers of greater stress and strain relative to sites of origination (Maas & Sandercock, 2010), (2) that some bones may be more or less robust to the mechanical environment than others (Ehrlich & Lanyon, 2002), and

DOI:10.6814/NCCU202100169

(3) that mechanical stimulation of bone development via muscle may act both locally, at a process, and globally, over the entire craniofacial skeleton (Chen et al., 2010; Schulte et al., 2013; Yucesoy, 2010)].

[NSS INT RA26]

In this example, three *that*-clauses coordinated by the conjunction *and* as well as numbers in brackets (e.g., (1)), along with its stance noun *predictions* and the inverting *colon*, constitute the *Noun Complement* construction. The colon here is also used to signal the forthcoming detailed explanation of the stance noun *predictions* in the cluster of *that*-clauses that immediately follows it. These two examples, taken together, seem to suggest that international writers in hard fields maintain a more reader-oriented attitude than their local counterparts, an attitude evident in their 46.7% of the overall punctuation mark use with colons, as compared to none of that in the local subcorpus. This can be attributed to the notion of "reader responsibility" in the East, where readers are made "primarily responsible for effective communication" (Hinds, 2001: p. 65), thus lifting writers' responsibility to guide readers through the text. Alternatively, it is also probable that due to local writers' unwillingness to draw on *that*-clauses of large grammatical complexity, there is thus no need for them to put a premium on the use of colons as post-modification.

Moreover, comma, as mentioned above, functions to connect two nominal structures, with the latter marked as the apposition of the former. It is interesting to note that although one may argue by intuition that comma seldom divides a stance noun and its complement clause as two nominal structures, the large proportions of commas are devoted to this pattern of use in both subcorpora, particularly in the international one, where Example (40) below is typically found:

(40) Physicists' nearly universal conviction**,** [that science should be formulated in a way that makes no reference whatever to the personal experience of the individual user of science, makes it impossible to express any of the above].

[NSS INT RA27]

DOI:10.6814/NCCU202100169

According to Biber and Gray (2016), academic writers since the 20th century have begun to draw on lengthy and complex appositive noun phrases to convey messages with a broad range of co-referential meaning relationships to the stance noun:

> *or still more strictly in relation to* ***Favus****, a disease dependent on the existence of a plant or vegetable like parasite-phyticides* (p. 205, italics and highlighting in the original)

Hence, *that*-complement clauses well suit these writers' need to substitute for those long and complicated appositive noun phrases, since clauses of this kind function as a noun but offer the space of a clausal structure, or in other words, this structure provides the most grammatical positions (e.g., verbal predicate) to include as many meaning elements as possible, in a way that makes easier the delivery of information of considerable length and complexity (Biber & Gray, 2016). The same holds true for natural scientists in both international and local journals, and their replacement of appositive noun phrases with *that*-complement clauses thus explains why the deployment of commas as post-modification constitutes a high proportion of the total punctuation use both internationally and locally. This is further proved not only by the use of commas in pairs (preceding and following the *that*-complement clause, thus indicating that it is a case of apposition), but also by no extra-linguistic effort made by authors through their comma use identified in the two subcorpora. Or to put it differently, comma is employed in this way out of pure grammatical purposes for the substitution of appositive noun phrases.

Finally, the function of dashes is similar to that of commas but associated with higher degrees of informality, which, thus, explains the over three times denser use of commas as post-modification than dashes in the international subcorpus. Namely, strict formality in academic writing urges authors to utilize more formal lexico-grammatical forms to convey information, such as commas. However, some researchers (e.g., , 2006; Swales et al., 1998; van Maanen, 1995) argue that trends toward increasing informality in scholarly writing has been emerging, "certainly abetted by post-modern and feminist movements, in which the scholarly voice is deliberately disassociated from detachment and distance" (van Maanen,

DOI:10.6814/NCCU202100169

1995, as cited in Swales et al., 1998: p. 118). Although the influence of post-modern and feminist movements over the hard fields is likely to be much less substantial than that in the soft ones, it is still reasonable to assume that this so-called "informalisation trend" (Ädel, 2006: p. 146), to a certain extent, "abets" international writers in the natural sciences to bring in academic written discourse some informal style of writing, through, for example, the dash use as post-modification. The 13.3% of its use by international natural scientists among all their use of punctuation seems to support this assumption, while this trend appears to be more noticeable in the local subcorpus, where the proportion rises to 50%. Nevertheless, this may be due to local authors' rhetorical unawareness of the informality of this pattern of use, given that dashes are not normally used to convey informal senses in the Chinese language, as in this example from the local natural science subcorpus:

(41) From the test results, the exciting possibility — [that the anomaly of quasi-static atmospheric electric field may really become a reliable mark for making short-term and imminent earthquake predictions] — is discussed.

[NSS LCL RA44]

5.2.3 Adverb phrase

The lexico-grammatical feature *adverb phrase* in post-modification use is most frequently presented by its linguistic role as code glosses in both international and local subcorpora, although there are merely three occurrences of such use by local authors (see Table 5.6 below). Code glosses usually assume the role of signaling or giving cues to the supplementary information offered through the explanation, elaboration, and rephrasing of the preceding text (Vande Kopple, 1985; Crismore et al., 1993; Hyland, 1998a, 1998b, 2004). As noted above, they reflect the writer's predictions about the reader's knowledgebase and can take the form of adverb phrases, such as *namely*, as in the following example:

(42) Recent advancements in singlet fission have been materials-limited due to the rarity of molecules that meet the essential energetic processual requirement,

DOI:10.6814/NCCU202100169

	NSS in CIJA	NSS in CCLJA
Endophoric marker	9 (16.1)	0 (0.0)
Code gloss	47 (83.9)	3 (100.0)
Totals	56 (100.0)	3 (100.0)

Table 5.6 Adverb phrase as post-modification between NSSs in CIJA and CCLJA (% of total)

namely, [that the energy of the lowest triplet excited state E(T_1) be on the order of half the energy of the lowest singlet excited state E(S_1)].

[NSS INT RA59]

In this case, the authors of this research article were likely to anticipate the actual content of the *requirement* going beyond the potential readers' existing knowledge, probably given that it is from "recent advancements" that they may fail to keep up to date with. Thus, the use of *namely* here not only discloses the authors' evaluation of the subject matter not shared by them and readers, but also hints at an authoritative authorial presence in comparison with readers (writers possessing more up-to-date information) (Hyland, 2004).

Example (43) is another typical case in the subcorpora, which, at the first glance, looks parallel to the last example:

(43) Indeed, the data are consistent with the opposite conclusion, **namely**, [that the reproducibility of psychological science is quite high].

[NSS INT RA13]

However, when placed in its co-text (Example (43) in **bold**), it seems to be open to an alternative interpretation:

A paper from the Open Science Collaboration (Research Articles, 28 August 2015, aac4716) attempting to replicate 100 published studies suggests that the reproducibility of psychological science is surprisingly low. We show that this article contains three statistical errors and provides no support for such a conclusion. **Indeed, the data are**

DOI:10.6814/NCCU202100169

consistent with the opposite conclusion, namely, that the reproducibility of psychological science is quite high.

In this excerpt, the authors first refer to the conclusion of a certain paper as erroneous and supportless, after which their own opposite conclusion derived from the same data is given, with the signaling adverb phrase *namely* providing the cue to its elaboration in the following *that*-complement clause. In effect, readers are supposed to know about the wrong conclusion from the very beginning of the text (i.e., "the reproducibility of psychological science is surprisingly *low*"), so there should have been no need for writers to reiterate the specific content of their own opposite conclusion, since it is just "the reproducibility of psychological science is surprisingly *high*". In this sense, the authors' use of the code gloss is grounded not in their assessment of readers' knowledgebase, or in indicating their authoritative position vis-à-vis readers, but rather in expanding a single noun *conclusion* into a long and specified *that*-complement clause to ensure its details and significance to be better taken in by readers.

Taken together, code glosses as post-modification in the *Noun Complement* construction, in general, are drawn on to signal the additional information in the complement of a stance noun based on how writers assess readers' existing knowledge by natural scientists in both subcorpora, or perhaps more importantly, to make extra-linguistic effort by just international writers, either to display authority or to ensure that the significance of their intended meanings is taken in by readers.

Owing to the single function of code glosses to signal additional information identified in the local subcorpus (see Table 5.6 above), it is not surprising that writers in domestic journals, compared to their international counterparts whose use of code glosses involves serving extra-linguistic purposes, seem to rely much less on it. Additionally, because this function of code glosses can be seen as a reader-friendliness feature of the text, local writers' lack of reader-oriented attitudes and their notion of reader-responsibility may, as demonstrated in their use of punctuation above (e.g., non-use of colons to guide readers), further explain why the employment of code glosses is neglected by them.

DOI:10.6814/NCCU202100169

Lastly, endophoric markers function to refer stance nouns to other parts of the text, which amounts to almost one fifth of all uses of adverb phrases as post-modification in the international subcorpus (see Table 5.6 above). Examples (44) and (45) are typical:

(44) Since $f_{n+1,a} = 0$, then $\alpha_{n+1} = 0$, and it follows from the analysis **above** [that the general equilibrium effect will be $dY = cd\tau_{n+1}$].

[NSS INT RA8]

(45) Given the size of the error bars, for this qualitative comparison the **warning** below [not to mix results at different scales, renomalization schemes, and order of perturbative expansion in α_s] has been ignored.

[NSS INT RA21]

According to McCarthy (2000), the most characteristic function of anaphoric expressions is to help establish the grammatical cohesion among individual clauses and utterances, or in other words, the cohesive organization of the overall discourse, which, in turn, contributes to the creation of "textuality", namely, "the feeling that something is a text, and not just a random collection of sentences" (p. 35). In this sense, the use of the endophoric marker *above* in Example (44) thus becomes a contributing factor in the establishment of grammatical cohesion and textuality of the specific research article where it is used. The endophoric marker *below* in Example (45), on the other hand, serves to cataphorically refer the stance noun *warning* to the subsequent text. A cataphoric expression is often characterized by its function of "engag[ing] and hold[ing] the reader's attention with a 'read on and find out' message" (McCarthy, 2000: p. 42). Hence, readers' attention in this case is also engaged and held by the endophoric marker *below*. It seems like making an announcement to all the readers that "if you are interested in knowing about more details surrounding this *warning*, you may read on and find out yourself".

As such, the use of endophoric markers aims for either contributing to the grammatical cohesion and textuality of the research article, or attracting and keeping the reader's attention to the following text. Unfortunately, these two useful linguistic and extra-linguistic functions are totally ignored by local writers in the natural sciences, as evident

DOI:10.6814/NCCU202100169

in none of its use in the local subcorpus (see Table 5.6 above). This is probably because they have far less awareness of the concepts of grammatical cohesion and textuality in English academic writing, while also refraining themselves from making such purposeful and motivated effort (e.g., engaging and holding readers' attention) as a way to rigidly observe the conventions of impersonality and objectivity in hard science articles. It is also reasonable to assume that they possess relatively lower levels of academic competence in English research article writing, since a significant aspect of academic competence of this kind is the ability to "establish a relationship with the reader" (Richards & Skelton, 1991: p. 34).

5.2.4 Extraposition

The use of *extraposition* as post-modification by international and local natural scientists, compared to that of prepositional phrases, punctuation and adverb phrases, exhibits the least statistical difference (r = -0.353, p < .001). Example (46) is typical in this pattern of use:

(46) But the simple fact **remains** [that narrative mosaics cannot really be assessed in any other way].

[NSS INT RA22]

In the above example, the "light", simple head noun *fact* is used in subject position, whereas its "heavy", complex *that*-complement clause is shifted toward the end of the sentence. In so doing, readers' comprehension of the sentence is eased, since they need not to carry the burden of retaining long and complex information from earlier in the sentence in their short-term memory, while also producing and understanding the remainder of it. Or, seen from a processing perspective, the *that*-complement clause is disruptive, because it must be processed before reaching the main clause predicate. Specifically, in the traditional grammatical sense, the core of a clause consists of two main parts: the subject and the predicate, which, together, convey the proposition (see, e.g., Miller & Miller, 2011; Paraskevas, 2020). As such, when the *that*-complement clause intervenes (especially in this case where the embedded clause is "heavy", complex), the proposition expressed by the

DOI:10.6814/NCCU202100169

joint effort of the subject and the predicate is then broken apart, thus hindering readers' understanding of the sentence meaning. Taken together, the use of extraposition, though causing the discontinuity within the *Noun Complement* construction, facilitates the reader's processing of the discourse through adhering to the information principle (a gradual increase in information load) and the principle of end-weight (heavy grammatical elements placed at the end of a clause).

However, one may argue that since a writer is not operating under severe production constraints, as compared to the speaker, who, given the real-time production constraints of conversation, must strictly observe the two principles in order for the hearer to decode the information. Thus, it is the writer's freedom to disregard such principles by leaving the *that*-complement clause in the position immediately after its head noun. In spite of this, most writers seldom violate these two principles if not for the discourse functions served by the pre-predicate *that*-clauses, the reasons of which are discussed as follows:

First of all, as Biber et al. (2007) noted, placing the "heavy" grammatical element before the main clause predicate (e.g., the placement of a *that*-complement clause right after its sentence-initial head noun) typically occurs, when the subject of the *that*-clause is a given referent, directly referring anaphorically to the previous text, while its predicate gives factual or generally acknowledged information about the referent. However, this does not well suit the situation of the *Noun Complement* construction, as the *that*-clause in such a construction is usually providing the value-laden information about the stance noun. Also, the subject of the *that*-clause is usually not a referent, given its role in providing the complete meaning of the head noun, thus not forming any anaphorical link to the preceding discourse, as *narrative mosaics* in Example (46) above.

Such a pre-predicate placement also prevails in the clauses beginning with *The fact that*, where long and complex elements, regardless of the information principle and the principle of end-weight, immediately follow the subject noun *fact*. This is done for the same reason that the complement is offering certain fact and generally accepted information, as the example below shows:

DOI:10.6814/NCCU202100169

> ***The fact that 14 of the 29 questions were answered correctly by 30 per cent or more of the lowest band*** *suggests that there is a range of questions within the conceptual grasp of all or practically all the lowest band of attainers*. (Biber et al., 1999: p. 678, italics and highlighting in the original)

Notwithstanding this, the cases where the stance noun *fact* is separated from its *that*-complement clause by the use of post-modification are rather rare in the two subcorpora, which is, thus, not able to account for local writers' reluctance to access the use of *extraposition*.

An additional factor affecting the inclination to use pre-predicate *that*-clauses instead of extraposed ones is the heavy use of long and complex grammatical elements following the main clause predicate, as in the example below:

> ***The example that a Lvf superiority in reaction time emerged at 100 msec.****, however, was taken [by Moscovitch et al.] [to mean that the right hemisphere had some advantage with respect to a more stable representation of the stimulus]*. (1999: p. 679, italics and highlighting in the original)

Here, the main clause verb *was taken* is followed by both a phrasal element *by Moscovitch et al.* and a complex *to-infinitive* complement clause "to mean that the right hemisphere had some advantage with respect to a more stable representation of the stimulus". In this respect, if *that*-complement clause were shifted to the end of the sentence, it would come after the other "heavy", complex elements, which, therefore, imposes a huge short-term memory burden on the readers. This is because they need to process all intervening elements before finally reaching the *that*-complement clause of the sentence-initial head noun *example*. Nevertheless, local natural scientists normally exhibit the general tendency for avoiding the use of "heavy", complex grammatical elements in their writing, as noted above (e.g., in the case of their non-use of colons). As such, their preference toward the use of a complex construction consisting of multiple phrases or clauses, such as that in the above example, is probably much slighter than that of their international counterparts. As such, their reluctance to increase this pattern of use also

DOI:10.6814/NCCU202100169

cannot be attributed to their dense use of long and complex elements following the main clause predicate which favors the pre-predicate over the extraposed *that*-clauses.

Last but not least, a writer's personal stylistic preference is likely to impact their employment of extraposition, because writers can, as noted above, simply opt to disregard and violate the general principle of shifting the "heavy" elements toward the end of the main clause. Hence, one may argue that the difference in this pattern of use is merely due to a few individual writers' stylistic preference for extraposition. However, this inferential statistical difference is in fact identified based on a series of one-to-one mathematical comparisons of the extraposition use between every single research article to each of the rest 299 articles in an overall database (NSS subcorpora) of 1,702,527 words (300 articles) via the nonparametric test *the Mann-Whitney U comparison*. Such a statistical test with a huge database basically counteracts the influence from individual authors' personal preferences. Moreover, in early literature, Clarke (1964) notes the trends of multiple authorship in science articles. Becher and Trowler (2001) also ascribe the characteristic of joint publication to "the hotter specialisms such as biochemistry" (p. 113). More recently, Tse (2012) reports the fast growth of multiple authorship in hard fields, with it even becoming a fairly common practice in certain disciplines such as engineering. All these findings are supported by the data from the present study that there are the averages of 5.4 and 6.7 authors per research article in international and local journals respectively. As such, an author's individual preference is further offset by the presence of other authors in each article. In addition, as Biber et al. (2007) suggest, the impact of individual preference is particularly noticeable with marked forms. However, the extraposition of *that*-clauses has the status as the unmarked form, given its higher occurrences than the pre-predicate *that*-clauses, which are, then, assumed to be the marked form.

Taken together, in general, the information principle and the principle of end-weight are firmly adhered to by writers, unless (a) the use of pre-predicate *that*-clauses serves discourse functions, (b) a "heavy", complex construction follows the main clause predicate, or (c) the personal stylistic preference comes into play. Nevertheless, the use of

DOI:10.6814/NCCU202100169

Feature	Mann-Whitney			Effect size (*r*)
	U	Z	*p*	
Punctuation	**3331.5**	**-10.681**	**< .001**	***r* = -0.617**
Adverb Phrase	**5947.5**	**-7.535**	**< .001**	***r* = -0.435**
Finite dependent clause	**7412.0**	**-7.048**	**< .001**	***r* = -0.407**
Non-finite clause	**7486.0**	**-6.031**	**< .001**	***r* = -0.348**

Table 5.7 Lexico-grammatical features of significant inferential statistical differences (in descending order of effect size *r* value) between SSSs in CIJA and CCLJA

extraposition in *Noun Complement* construction normally does not relate to these three scenarios. In other words, local natural scientists are hardly likely to deploy less extraposition because they are under any of these scenarios. Thus, their reluctance to access extraposition is most probably due to their lack of knowledge about the two principles.

5.3 Explanations for different use between international and local writers in the social sciences

Inferential statistically, the variances in the use of post-modification structures between international and local social scientists with the Pearson *r* value of at least ±.300 have been found in that of punctuation, adverb phrases, finite dependent clauses, and non-finite clauses (in descending order of effect size *r* value; see Table 5.7). Since the most statistically significant difference between these two groups of writers lies in their use of punctuation, the analyses in this section will start with this pattern of use.

5.3.1 Punctuation

The four forms of punctuation as post-modification have been identified from the two subcorpora as *comma*, *colon*, *dash* and *semi-colon*. Proportionately, comma tends to dominate the punctuation use as post-modification in the local subcorpus, while it also accounts for the largest percentage of such use in the international subcorpus (see Table 5.8 below). Although there is proportionately more dashes in the local subcorpus, the density of colons in the international one increases to over two times as high as that in its local counterpart, also with a proportion, though relatively small, of the punctuation use

DOI:10.6814/NCCU202100169

	SSS in CIJA	SSS in CCLJA
Comma (,)	379 (37.8)	58 (46.4)
Colon (:)	342 (34.1)	21 (16.8)
Dash (—)	269 (26.8)	46 (36.8)
Semi-colon (;)	13 (1.3)	0 (0.0)
Totals	1003 (100.0)	125 (100.0)

Table 5.8 Types of punctuation used between SSSs in CIJA and CCLJA (% of total)

represented by semi-colons.

Comma, as the dominant punctuation mark in this pattern of use, is usually used to introduce elaborating or explaining elements to the content of the head noun in front of it. With the content of apposition becoming increasingly long and complex as noted earlier, academic writers tend to substitute the nominal complement clause for the appositive noun phrase. This is due to its clausal structure, which, as compared to phrasal structures (e.g., noun phrase in this case), provides the most grammatical positions to include as many meaning elements as possible within one single syntactic structure. Or in Biber and Gray's (2016) terms, clausal forms "are considerably more explicit" than phrasal ones, granted that they "grammatically specify the meaning relationships among elements" (p. 18). Example (47) is typical in both subcorpora.

(47) I have tried as hard as I know how to have accepted this <u>idea</u>**,** [that where Federal funds and Federal authority are involved, there should be no discrimination based upon any reason that is not recognized by our Constitution].

[SSS INT RA6]

In the field of social sciences, the nominal complement clause separated by a comma from its head noun typically serves the function of demonstrating an identity meaning relationship with that noun. Specifically, such clauses are commonly used to clarify the identity of or give descriptive information to help specify the head noun (see Example (47)). Notwithstanding this, international social scientists incline to make strategic effort in their use of commas as post-modification as a way to extend the general function of it. As shown

DOI:10.6814/NCCU202100169

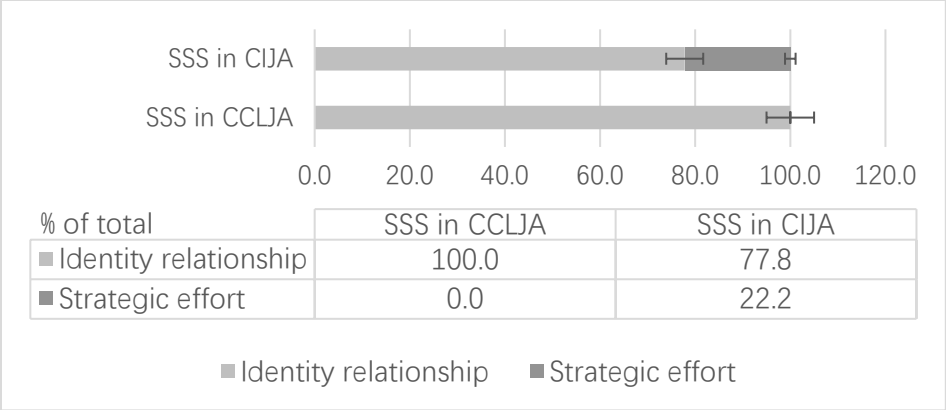


Figure 5.1 Types of *comma* use between SSSs in CIJA and CCLJA

in Figure 5.1, 22.2% of commas are used by international writers to make their strategic effort (for detailed explanations of this function, see Section 4.3 above).

It should be emphasized that because more than 20 percent of commas is devoted to making the strategic effort in the international subcorpus, if, proportionately, only the comma used with the identity meaning relationship taken into account, the difference in the percentage of such use between international and local subcorpora is even larger than that demonstrated in Table 5.8 above. Specifically, when considering the *comma* use for the identity meaning relationship alone, it only accounts for 29.4%[2] of the overall use of punctuation as post-modification in the international subcorpus, as compared to 46.4% in its local counterpart. One might argue that this is due to local writers' tendency to express more identity relationships than international writers. Nevertheless, according to Biber and Gray (2016), the function of expressing an identity meaning relationship can be served by other lexico-grammatical features apart from nominal complement clauses, such noun phrases, adjective phrases, as well as non-finite and finite clauses of various kinds. As such, local writers' more frequent use of commas cannot be simply attributed to their inclination toward the expression of more identity relationships. Besides, it seems clear that local writers tend not to make full use of commas, given their disregard of its role in writers'

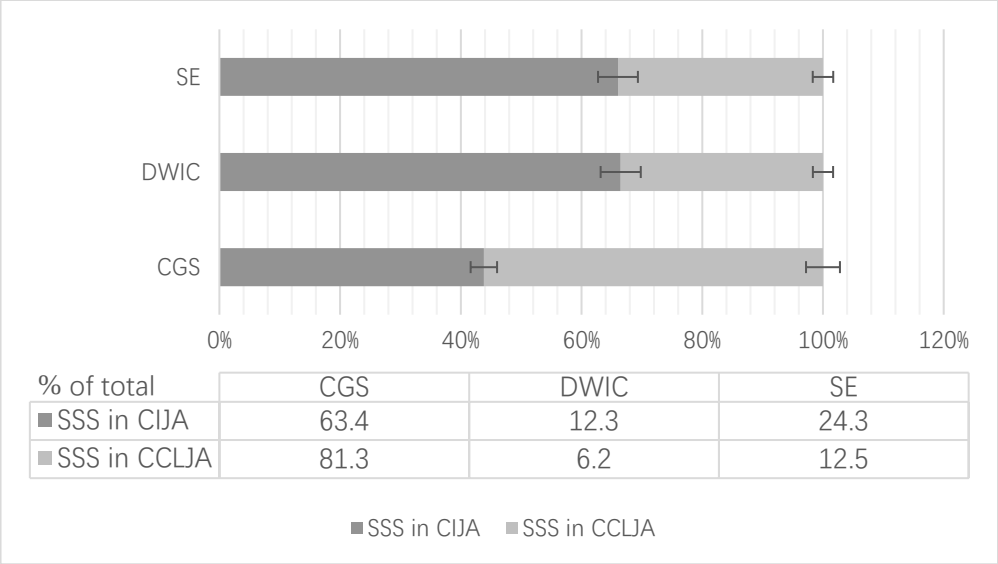
[2] 29.4% = 37.8% (percentage of the total *punctuation* use) × 77.8% (percentage of the total *comma* use)

DOI:10.6814/NCCU202100169

launch of strategic effort (see Figure 5.1 above). This might be caused by their lack of enough awareness of how post-modification structures can be used as "epistemic strategies" (Holms, 1988: p. 32) in the genre of research articles. Besides, as one significant aspect of academic competence is the ability to establish a relationship with the reader, the competence to strategically bring readers into alignment with writers' points of view is obviously critical to such a relationship. However, local writers' relatively lower levels of academic competence in this regard prevent them from doing so. As such, whenever they choose a comma to separate a head noun and its nominal complement clause, they have no choice but to use it only for the identity relationship function, the proportion of which, thus, grows in the local subcorpus.

Another feature that sets local writers apart from their international counterparts is the density of colons, which is, as shown in Table 5.8 above, less than half as high in the local subcorpus as in the international one (16.8 vs. 34.1). When placed in a *Noun Complement* construction, where a colon functions to connect a head noun and its nominal complement clause, it signals the meaning relationship between its preceding word/phrase and following clause, namely that the latter specifies or explains the content of the former. In most cases, the following clauses are, as demonstrated in the natural sciences subcorpora, of great grammatical complexity, such as in the form of coordinated *that*-clauses. This finding is supported by the most frequent use of colons to connect the head noun and the complex grammatical structures (i.e., CGS) as its complement in the social sciences subcorpora as well (see Figure 5.2). Despite local natural science writers' ignorance about this pattern of use, their social science counterparts draw on even more of such a construction than their international counterparts (81.3 vs. 63.4). More importantly, social scientists, unlike their natural counterparts, extend the range of the colon use to include linking the head noun and its dependent *wh*-interrogative clause (i.e., DWIC), as well as the head noun expressing the substantive meaning and its complement (i.e., SE). In general, the use of colons for DWIC and SE in the local subcorpus of social sciences is merely over half as frequent as that in the international one. Specifically, in the case of SE, although

DOI:10.6814/NCCU202100169



% of total	CGS	DWIC	SE
SSS in CIJA	63.4	12.3	24.3
SSS in CCLJA	81.3	6.2	12.5

Note: CGS = Complex grammatical structures, DWIC = Dependent *wh*-interrogative clauses, SE = Substantive expression.

Figure 5.2 Types of *colon* use between SSSs in CIJA and CCLJA

the grammatical elements placed in the complement position of the head noun are not long and complex, as illustrated in Example (48), a colon is still drawn on to connect the head noun and its complement, given the substantiveness of the head noun essential to writers' conveyance of the core messages in their work.

(48) Columns (4)–(6) compare *establishment* and *hierarchy* and draw the same conclusion**:** [that the effect is primarily evident in the early stage].

[SSS LCL RA13]

Here, the head noun *conclusion* belongs to the category of substantive expressions, for it is reasonable to assume that this "same" *conclusion* plays an essential role in the authors' argumentation; otherwise, they need not to reiterate it ("same" implies that the conclusion has been presented before in the text), and to particularly use a colon to link it with its *that*-complement clause. In other words, although the information contained in the *conclusion* is given rather than new, it is still essential to the writer's construction of the intended meaning for the readers, namely that the conclusion is vital so do not miss this main point.

DOI:10.6814/NCCU202100169

Thus, its specified content in the following complement clause is introduced by a colon to highlight its substantiveness.

The reason why social scientists choose colons rather than other punctuation marks for their substantive expression use is perhaps because colon, as Huddleston and Pullum (2002) put it, is a stronger mark of formal style than comma, while dash is "the least formal" mark (Seely, 2007), which, thus, prioritizes the use of colons, given the high level of stylistic formality of academic writing. As such, it seems reasonable to predict that local writers' reluctance to employ more colons for the use of substantive expressions is mainly because they are likely to give little heed to the substantiveness of different head nouns that they use, whereby such nouns with substantiveness might be treated like the common ones, thus rendering the use of colons as a connector unnecessary. From another perspective, this also proves their unfamiliarity with the rhetorical conventions in English research article writing which entail writers' full construction and expression of the intended meanings for their readers through different lexical-grammatical items, even a colon as in this case.

Moreover, the placement of dependent *wh*-interrogative clauses in complement position after the colon accounts for the smallest proportion of the total use of colons as post-modification by both international and local writers (see Figure 5.2 above). As demonstrated in the following example, the dependent *wh*-interrogative clause in complement position is not that long or complex as, for example, coordinated *that*-clauses. In fact, it is just a finite dependent clause, marked by a *wh*-word as the clause link. Moreover, compared to the use of substantive expressions, the "substantiveness" of the head noun in the dependent wh-interrogative clause is not that salient. To illustrate, the head noun *variable*, in Example (49), is simply a series of statistical numbers mentioned by the authors that they have substituted for the numbers of the "five-year uptake indicator":

(49) The regression results we present in Columns 4 and 5 also analyze expungement receipt, but they replace the five-year uptake indicator with a different outcome variable**:** [whether an expungement is received in any particular quarter].

[SSS INT RA72]

DOI:10.6814/NCCU202100169

This is further confirmed in the co-text of this example, where the core messages of the entire section are the results of the statistical analyses rather than a single variable in those analyses (Example (49) **in bold**):

> **The regression results we present in Columns 4 and 5 also analyze expungement receipt, but they replace the five-year uptake indicator with a different outcome variable: whether an expungement is received in any particular quarter.** These analyses help us to understand the influences on expungement probability that vary over time — and in particular, to ask whether an individual's immediate employment history (such as a recent job loss) drives the decision to apply for an expungement.

Example (50) is also typical:

(50) The current controversy over universal injunctions against federal law has been framed around one simple question**:** [whether a federal court can "prohibit the Government from enforcing a policy with respect to anyone, including nonparties]."[239]

[SSS INT RA41]

The author, in this example, describes the head noun *question* as "simple". Actually, all the words here are used to deliver the background information instead of the core message. According to Huddleston and Pullum (2002), colon is sometimes taken to mark the boundary between the declarative and the interrogative sentence, with it followed by a capital letter in some cases:

i *The same question occurred to both of them: why had no one called the police?*

ii *A number of questions remain to be answered: Who will take responsibility for converting the records to digital form? How are the old records to be stored? Who will have access to the digital files?* (p. 1736, italics in the original)

In this sense, dependent *wh*-interrogative clauses, to a certain extent, closely resemble interrogative sentences. This is largely supported by the fact that the full stop at the end of

DOI:10.6814/NCCU202100169

the dependent *whether*-clause in Example (50) can be simply replaced by a question mark, without changing its meaning:

> The current controversy over universal injunctions against federal law has been framed around one simple question: whether a federal court can "prohibit the Government from enforcing a policy with respect to anyone, including nonparties**?**"[239]

As a result, the reason why a single dependent *wh*-interrogative clause, without the complex grammatical structure and being essential to authors' argumentation, is still connected by a colon with its head noun might be due to its obvious resemblance to an interrogative sentence, and is thus used in a nearly parallel way. Therefore, local writers' little exposure to this particular pattern of use is likely to generate their uncertainty about whether it is grammatically appropriate to place a single *wh*-clause in the complement position, introduced by a colon, of a head noun.

Lastly, local writers, on the contrary, make more use of colons to connect a head noun and the complex grammatical structures in its complement position. This use, as noted above, displays the "reader-friendliness" of a text. One may argue that this demonstrates that local social scientists maintain a more reader-oriented attitude than their international counterparts. Nevertheless, this may also be due to their indiscriminate use of colons, commas, and dashes, irrespective of their patterns of use and levels of formality, which might, then, result in the placement of colons in the positions where there should have been commas or dashes. This is further proved by one distinctive feature that sets the punctuation system in English apart from that in the Chinese language: the ubiquity of colons in modern Chinese as "the express signals of quoted speech and well-planned written discourse" (Zhang, 2017: p. 65). Hence, the tendency of local writers as Chinese native users to draw on more colons, along with their grammatical and rhetorical unawareness that colons can be deployed either with the substantive head noun or the dependent *wh*-interrogative clause (which leaves them no choice but to use colons only for introducing complex grammatical structures), contributes jointly to their denser use of colons to link head nouns and complex elements in their complement positions.

DOI:10.6814/NCCU202100169

Thirdly, local writers more favor the use of dashes as a way to introduce grammatical elements that further develop or exemplify the head nouns than their international counterparts (36.8 vs. 26.8). As Huddleston and Pullum (2002) note, in terms of attaching a head noun to its nominal complement clause, comma and dash are often in competition with each other. Although dash seems to mark a noticeably stronger break from the immediate co-text than comma, while also permitting a wider variety of grammatical elements (e.g., a main clause) to follow it, comma remains at a higher level of formality than dash, which is even seen as "the least formal" mark. In fact, for the use of punctuation marks to connect a head noun and its complement in the *Noun Complement* construction, the grammatical element following a punctuation mark is fixed, namely, nominal complement clause, while the break from the surrounding text is also underplayed under such circumstances, since the complement clause, though separated from its head noun, relates closely to it, as demonstrated by the co-referring relationship between them, as well as the elaboration of the head noun by the complement clause. Therefore, the levels of formality between comma and dash instead of the two merits of the latter are thus prioritized in scholarly writing, or in other words, academic writers are more liable to make more use of the punctuation mark with the higher level of formality.

However, the influence of post-modern and feminist movements, as discussed earlier, contributes to the trend toward increasing informality or "informalisation" in academic writing, which, to a certain extent, may cause academic writers' increased reliance on "the least formal" dash in their post-modification use. Notwithstanding this, such an influence is on international and local writers alike, so it is clearly not possible to argue that local writers have been affected more by post-modern and feminist movements than their international counterparts; actually, the opposite may be more probable, given the widespread popularity of these movements in English-speaking countries. More importantly, in the Chinese language, it is not uncommon for dashes and colons to be interchangeably used (Li, 2018), which might, then, lead to the effect of negative transfer (see, e.g., Benson, 2002; Ellis, 2008) on local writers' awareness of the levels of formality

DOI:10.6814/NCCU202100169

among punctuation marks, thus causing them to treat dash and colon as being at the same formality level. Furthermore, Chinese writers are more liable to replace colons with dashes, but not vice versa. This is because in the Chinese language, dash often introduces something which further develops or exemplifies what has been written before, which can be omitted from a text without changing its meaning, thus rendering its use more grammatically flexible (e.g., being easier to insert it into a clause), and it also serves the function of emphasizing the text introduced, whereas in general, colon just helps list items which are rather lengthy and therefore cannot be deleted from the text (see, e.g., Lin, 2000; Zhong, 2006).

Taken together, it seems reasonable to assume that local writers' misuse of dashes to function as colons, together with their ignorance of informality levels among punctuation marks, may give rise to their stronger preference for the use of dashes to connect a head noun and its complement than their international counterparts.

Last but not least, in terms of semicolons as post-modification, only the international subcorpus contains its infrequent occurrences. According to Huddleston and Pullum (2002), semicolon allows elaborative interpretations, which means that what has been written immediately following a semicolon can be used to elaborate or interpret something that precedes it. Sanchez-Stockhammer (2016) further pointed out that semicolon, when used for elaborative interpretations, usually marks a break between two parts of a sentence so as to display the close connection between them. Example (51) is typical in the international subcorpus:

(51) The results presented so far have assumed a symmetric treatment effect**;** [that banks respond to tax increases (cuts) by reallocating lending towards (away from) other states where they lend].

[SSS INT RA29]

To serve this function of elaborately interpreting something preceding it, semicolon is actually in competition with the punctuation marks *comma* and *dash*; however, in a relatively formal style, such as in the research article as in this case, writers typically choose

DOI:10.6814/NCCU202100169

a semicolon rather than a comma, needless to say a dash (Huddleston and Pullum, 2002). Sanchez-Stockhammer (2016) even found the occurrences of the semicolon only in scholarly writing. More importantly, academic writers are, according to Biber and Gray (2016), often motivated to draw on a semicolon due to the great length and complexity of what has been written that follows it. As in Example (51), although the nominal complement clause contains only one main clause, it actually combines one preposition with a non-finite clause as its object, two appositive phrases in parentheses and an embedded relative clause with the main clause. One may argue, at this point, that in this function the semicolon could be replaced by a colon. Nevertheless, colon could occur as a less preferred option in these cases, since it is more commonly followed by longer and more complex grammatical elements (e.g., a list of advice points, each taking the form of a *that*-complement clause) than semicolon.

Additionally, despite the fact that semicolon, in its function of elaborative interpretations, tends to mark a break between two closely related parts of a sentence, these two clauses are, as Huddleston and Pullum (2002) and Sanchez-Stockhammer (2016) note, normally independent finite clauses:

> *The bill was withdrawn; the sponsors felt there was not sufficient support to pass it this session.* (Huddleston and Pullum, 2002: p. 1742, italics in the original)

This, hence, explains why the use of semicolons only occupies such a small proportion of all punctuation use as post-modification: namely, that the nominal complement clause separated by a semicolon from its head noun belongs to the type of dependent rather than independent finite clauses, which, in turn, reduces its occurrences after semicolons. Meanwhile, given such a grammatical constraint on the structural type of clauses that follow a semicolon, it is thus not surprising why local writers make no use of such a construction.

5.3.2 Adverb phrase

In the field of social sciences, writers, in comparison with their counterparts in hard fields, tend to draw on adverb phrases to express a broader range of linguistic and extra-

DOI:10.6814/NCCU202100169

linguistic functions (see Table 5.9). According to Vande Kopple (1985), Crismore et al. (1993), Hyland (1998a, b, 1999b, 2004, 2005b) and Hyland and Tse (2004), such functions can be subsumed under two general headings of *textual* (or in what Hyland (2005b) calls *interactive*) and *interpersonal metadiscourse*. The present study follows Hyland's

	SSS in CIJA	**SSS in CCLJA**
Textual (*Interactive*)	*308 (77.5)*	*56 (94.9)*
Transition markers	63 (15.8)	13 (22.0)
Frame markers	39 (9.8)	7 (11.9)
Endophoric markers	2 (0.5)	2 (3.4)
Code glosses	204 (51.4)	34 (57.6)
Interpersonal	*89 (22.5)*	*3 (5.1)*
Hedges	28 (7.1)	3 (5.1)
Boosters	31 (7.8)	0 (0.0)
Attitude markers	30 (7.6)	0 (0.0)
Totals	397 (100.0)	59 (100.0)

Table 5.9 Adverb phrase as post-modification between SSSs in CIJA and CCLJA (% of total)

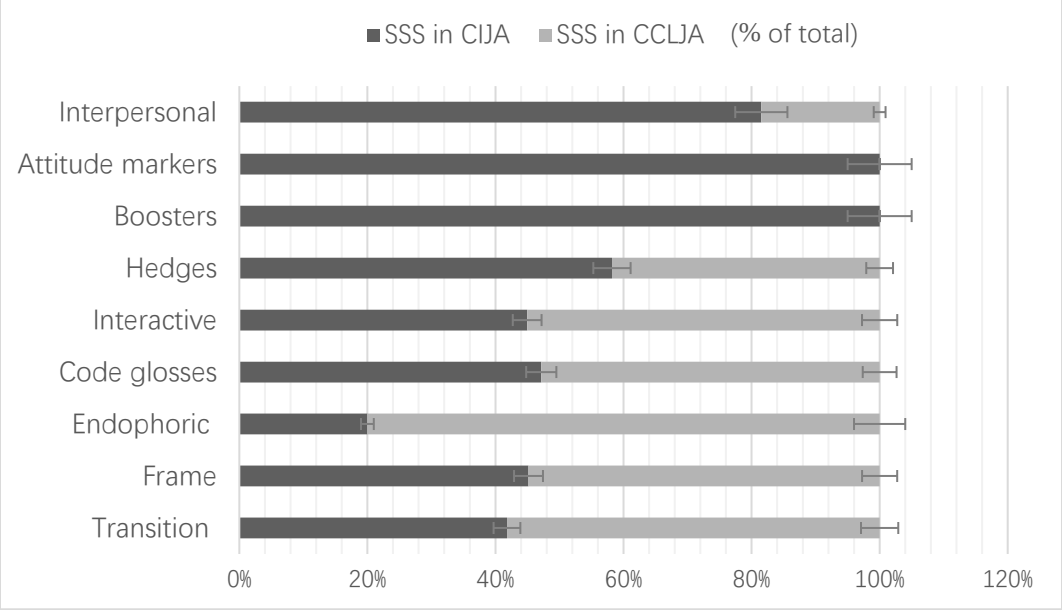


Figure 5.3 Adverb phrase as post-modification between SSSs in CIJA and CCLJA

DOI:10.6814/NCCU202100169

substantial revisions of Vande Kopple's and Crismore et al.'s classification of these two types of metadiscourse, which, to a large extent, collapse, separate and reorganize their categories. He thus widens the scope of, in particular, the interpersonal metadiscourse strategies, with the aim of focusing more on the interactions between writer and reader and the social functions of metadiscourse. As such, these functions are mainly concerned with the interpersonal relationship between writer and reader.

In general, the international subcorpus includes a significantly higher frequency of adverb phrases as post-modification, compared to that of its local equivalence, for almost each type of metadiscourse strategies (except code glosses for which the two are equal). Notwithstanding this, proportionally, local writers favor the use of interactive metadiscourse more than their international counterparts, while the latter's preference for interpersonal metadiscourse is much stronger than the former (see Figure 5.3).

Specifically, the largest difference in interactive forms between international and local writers lies in their use of adverb phrases to function as endophoric markers. Endophoric markers, as noted above, are expressions used to refer readers elsewhere in the text for additional information (e.g., *above*, *below*), which, thus, not only facilitates their understanding of the author's intended meaning, but also lends support to the author's argument through the above-mentioned and forthcoming material in its co-text, as illustrated in Example (52).

(52) There is clearly a need for an objective evidence base for its application (hence our <u>suggestion</u> **above** [of applying economic data and data on the demand for public services]).

[SSS INT RA49]

Here, the use of the adverb *above* separates the head noun *suggestion* from its *of* + *ing*-complement clause "of applying economic data and data on the demand for public services". The reiteration of the *suggestion* mentioned in the previous text seems to reflect the author's assumption about the reader's processing need at this point of reading this article, namely, that the reader is likely to forget or merely have a vague recollection of the

DOI:10.6814/NCCU202100169

specific content of the *suggestion*, given the fact that the first mention of it is 14 pages back in the article (with 73 pages in total from a law journal). In so doing, the reader's impression can be reinforced, thus helping them have a better comprehension of the author's argumentation here, since this above-mentioned *suggestion* also provides part of the justification to support "a need for an objective evidence", as can be seen in this example. Example (53) is typical in the two subcorpora as well:

(53) That culture is seen to overlap with the boundaries of a society and 'imagined community' (the term originally applied by Anderson (1991) in defining Nationality, also used by Tamir, 1993, 1995) supports my claim **below** [that the raw materials behind the definitions of nationality exist in other collectives].

[SSS INT RA73]

Given these two rhetorical functions performed by endophoric markers, one may argue that the larger proportion of adverb phrases used as such markers in the local subcorpus indicates that local writers take more care of the readers' need for the aid in comprehending the research article and possess a higher level of ability in the building of strong arguments at the discoursal level. Nevertheless, if the use of endophoric markers by writers in the field of natural sciences is also taken into account, a clearer picture of why international social scientists make the least use of this marker is likely to emerge. Specifically, nearly one fifth of all metadiscourse devices is employed as endophoric markers in the international subcorpus of natural sciences, although local natural scientists, given their disregard of grammatical cohesion and textuality and of engaging and holding readers' attention, make no use of such devices. It is not surprising that writers in hard fields, in light of their stricter adherence to impersonality and objectivity in academic writing (Charles, 2006; Gross & Chesley, 2012; Hyland, 2004, 2005b, 2009, 2012; Myers, 1990), tend to display a certain reluctance to access endophoric markers, granted that greater impersonality and objectivity can be achieved through few reflexive references to the text or little help to the reader in grasping the intended meaning of the author (Hyland, 2007; Mulholland, 1999). Notwithstanding this, endophoric markers still account for a greater proportion of all

DOI:10.6814/NCCU202100169

metadiscourse use in the international subcorpus of natural sciences than any of the two social science subcorpora. Example (54) is typical in the international natural science subcorpus:

(54) This confirms our assumption **above** [that our activation energy of diffusion of 4.4 eV is the sum of a defect formation and migration part].

[NSS INT RA15]

So, the question naturally arises as to why international social scientists, with relatively loose adherence to the conventions of academic writing in the soft knowledge domains, are less liable to draw on reflexive references and aid readers to understand their intended meanings through the use of adverb phrases as post-modification in the *Noun Complement* construction?

According to Becher and Trowler (2001), soft knowledge field has at its core the quest for the comprehension of cognitive entities, processes and events, and thus the construction of theories and arguments about such understanding. As a result, writers in soft domains tend to construct arguments and knowledge based on their personal interpretations and negotiations with readers, which are, in turn, open to the evaluation and judgement of writers (Charles, 2007; Hyland, 2005b, 2014). Nevertheless, new arguments and knowledge are not advanced merely from writers' own thoughts and ideas but rather "follow[] altogether more reiterative and recursive routes as writers retrace others' steps and revisit previously explored features of a broad landscape" (Hyland, 2004: p. 31). Similarly, such author-centered epistemic and evaluative judgment is also formed based on the evaluation of both their own and others' work (Charles, 2007; Hyland, 2005a). As such, for social scientists, as compared to their natural counterparts, the research article is more of "a multilayered hybrid *co-produced* by the authors *and* by members of the audience to which it is directed" (Knorr-Cetina, 1981: p. 106, original italics).

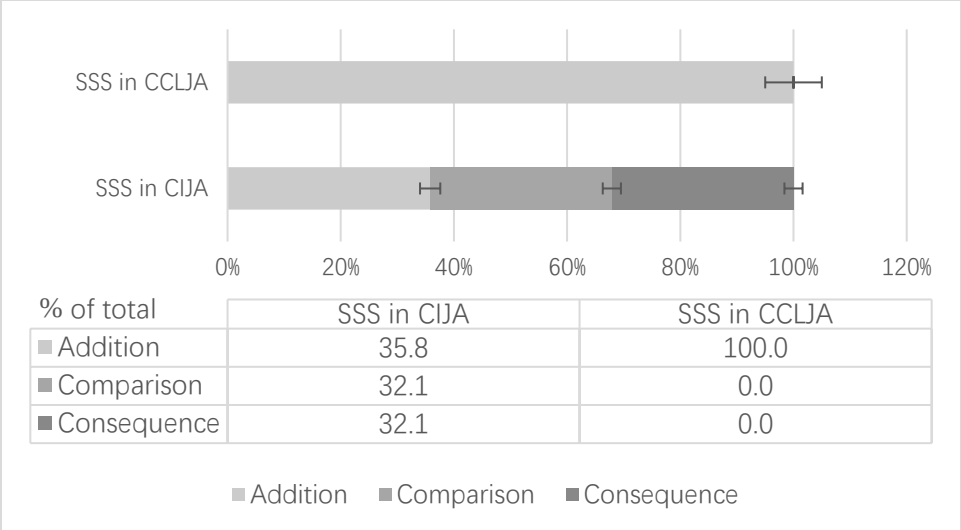
In this sense, it is reasonable to assume that writers in the soft knowledge fields are perhaps keener to convey an interpersonally acceptable persona to members of their disciplinary communities, by, for example, attending to the affective expectations of

DOI:10.6814/NCCU202100169

readers, showing proper respect for them, and even, in Myers's (1989) terms, "present[ing] themselves as equally the humble servants of the discipline" (p. 4). Thus, the bare use of a single adverb, such *above* or *below*, as a way to direct readers to other parts of the research article may seem less polite, given its impersonality suggesting the distance between writers and their readers, and more importantly, its directive- or
imperative-like form (like saying that "go find it up/down there yourself"). According to Hyland (2005b), directives or imperatives are viewed as risky in the soft fields where writers tend to treat them with caution. Therefore, the reluctance of social scientists to employ the form is related to its impersonality, directive/imperativeness and covert impoliteness, which might be, to a large extent, disregarded by local writers in the soft domains, given their lack of access to the above English academic community and genre conventions, thus contributing to their overuse of it to direct readers through the text.

Proportionally, the level of difference in the use of transition markers between international and local writers in the social sciences follows that of endophoric markers. As Hyland (2005b) suggests, transition markers are commonly employed to show relations between main clauses, which, then, can be further divided into three sub-types: namely, *addition* for adding arguments, *comparison* to compare and contrast arguments, and *consequence* relations either informing readers of conclusions being drawn or of arguments being rebutted. As can be seen in Figure 5.4, although there is proportionately less use of transition markers by international writers than their local counterparts, the latter draw exclusively on additions to help readers identify the addictive connections between arguments, as compared to the former, whose use of this metadiscourse marker is almost evenly devoted to all of its three types (see Examples 55, 56 and 57 below). Transition markers, as noted above, function to render the text structure of a research article explicit. Therefore, two plausible scenarios are developed here to account for the difference in the use of transitions between international and local writers: firstly that the argument structure of the article in local journals is relatively simple, void of lengthy and complicated explanations, justifications and discussions, so that its author simply needs to employ the

DOI:10.6814/NCCU202100169



% of total	SSS in CIJA	SSS in CCLJA
Addition	35.8	100.0
Comparison	32.1	0.0
Consequence	32.1	0.0

Figure 5.4 Adverb phrase as transition markers between SSSs in CIJA and CCLJA

(55) Under the assumption that the court correctly sets due care equal to the socially optimal level of care (=x*),[138] and **also** [that the court sets the magnitude of compensatory liability (L) equal to the actual harm (H), then the total accident costs faced by a potential injurer is a piecewise function expressed as follows:

$$T(x) = \begin{cases} C(x) + p(x)\mathrm{L}, if\ x < \tilde{x} \\ C(x), if\ x \geq \tilde{x} \end{cases}$$

where $x = x^*$ *and* $L = H$ ($Expression$ 2).]

[NSS LCL RA45]

(56) Discussions of nonbinary gender rights are often stifled by the assumption that those rights must always take the form of gender neutrality or, **alternatively**, [that the law must always recognize a third gender].

[NSS INT RA41]

(57) They attached great importance to the idea that a solution, when found, would apply to all games in (at least) a very broad class, and **therefore** [that an important property of prospective solutions should be that they should exist for all games].

[NSS INT RA28]

DOI:10.6814/NCCU202100169

interactive form of metadiscourse *addition* to add the arguments together, and secondly, that local writers' arguments are of equivalently great complexity as that of their international counterparts, but the exclusive use of additions in their articles is unlikely to guarantee the clearness and comprehensibility of the long and complex content to readers and thus to aid them in grasping the writer's intended meaning.

There is no denying that the first scenario is playing out in certain research articles, where the writers are striving to keep their arguments direct and simple; nevertheless, such cases are rare in the soft knowledge domains, since most of their findings are explained through argumentation and persuasion rather than experimentation and scientific observation. As such, the second scenario seems to be more probable, under which local writers' complete disregard of using comparisons and consequences (see Figure 5.4 above) may support the view that Chinese native writers, no matter whether they are student writers (e.g., Bolton et al., 2002; Lei, 2012) or expert writers (e.g., Gao, 2016), are all liable to significantly overuse phrase- and sentence-level additions as a way to compose unnecessarily lengthy and complex sentences. This can be attributed to local teachers' prescriptions for the demonstration of high language proficiency since students' early stage of learning a second/foreign language (Gardezi & Nesi, 2009). Alternatively, this may simply reflect their lack of familiarity with these two metadiscourse devices (i.e., comparisons and consequences).

Lastly, both international and local writers in the social sciences exhibit the similar preferences for the use of frame markers and code glosses. The metadiscourse device *code gloss*, as already pointed out, functions to introduce the elaboration of propositional meanings in the text (e.g., *namely*), while, according to Hyland's (2004, 2005b, 2009) terminology, frame marker is typically used for the signal of sequencing (e.g., *firstly*), the label of text stages (e.g., *overall*, *so far*), the announcement of discourse goals (e.g., *goal*, *aim*), or the indication of topic shifts (e.g., *now*, *right*). As noted above, the construction of arguments and knowledge in the social sciences is grounded in soft filed writers' personal interpretations and negotiations with readers, which, in turn, infuses the discursive nature

DOI:10.6814/NCCU202100169

into the lines of inquiry in these domains. Moreover, arguments and knowledge are usually constructed by following and revisiting others' work, and accessing the existing literature from a wide variety of academic, historical and topical territories (Bazerman, 1988; Becher, 1989; Becher & Trowler, 2001), which, in turn, encompasses the relationships and variables that tend to be larger in number, harder to delineate and more subject to contextual and interactional unpredictability (Hyland, 2014). This thus renders the argument structure in soft fields more diversified, reiterative and non-linear. As such, there has been a more urgent necessity for social scientists to guide readers through the text by virtue of any rhetorical devices, which, of course, include the metadiscourse strategies *frame marker* and *code gloss*. In so doing, writers are more likely to ensure that readers would not only not miss the main points of the text, but also not misread their intended meanings. On the part of local writers, more importantly, the use of frame markers and code glosses does not entail any complex and sophisticated syntactic structures or grammatical roles (cf. the case of *qualifier + subject* above), as illustrated in the following examples from the local social science subcorpus:

(58a) The results provided two indisputable facts, **firstly** [that SRS models are more accurate in certain flow situations than RANS models].

[SSS LCL RA8]

(58b) Therefore it should be required to undertake the so-called unreal joint liability, **namely** [that users may demand the franchiser or the franchisee to take all the responsibility, then the franchiser who has undertaken the compensating responsibility should have the right of recourse against the franchisee who is exactly at fault on the delivery process].

[SSS LCL RA39]

Here, in Examples (58a), the frame marker *firstly* is simply placed between the head noun *facts* and its *that*-complement clause, as a way to indicate the following complement being the first one in the sequence vis-à-vis all other facts to be unfolded in the subsequent texts. Similarly, the code gloss *namely*, in Example (58b), is employed to connect the head

DOI:10.6814/NCCU202100169

noun *liability* and its complement in a way that introduces the following *that*-complement clause as the elaboration of the head noun. Taken together, both syntactic structures and grammatical roles are rather simple in the use of these two interactive forms, which, along with the intrinsic urge for social scientists to guide readers through the text, perhaps contributes jointly to local writers' frequent employment of adverb phrases as frame markers and code glosses in their post-modification use.

The deployment of interpersonal metadiscourse, on the other hand, witnesses the distinct differences between international and local social scientists. In the local journals, writers express great reluctance to draw on the metadiscourse marker *hedge*, along with their total disregard for *booster* and *attitude marker*, whereas international writers exhibit almost equal preferences for all three devices (see Table 5.9 and Figure 5.3 above). Hedges (e.g., *almost*, *generally*, *perhaps*) are, in early literature, simply referred to as "words or phrases whose job it is to make things fuzzier" (Lakoff, 1972: p. 195). Zuck and Zuck (1986) further define it as the process by which the writer attenuates the strength of a statement, while for Markkanen and Schröder (1989) and Skelton (1997), it can be any manipulation of indirect communicative strategies to state less than one fully intends to mean. More recently, a clearer and more comprehensive definition of hedge is proposed, according to which it is the communicative device for the avoidance both of giving a clear and complete commitment to the truth of propositions on the one hand, and of exaggerating what is written and thus incurring the negation from readers on the other. In another sense, it allows writers to advance a proposition as an opinion but not a fact, which is, in turn, open to negotiation with readers, displaying writers' willingness to acknowledge standpoints and voices from other members in the disciplinary communities (Holmes, 1995; Hyland, 2005b, 2014; Nash, 1990; White, 2003).

There seems to be a general consensus among scholars and researchers that hedge is a vital piece in the construction of academic written discourse, granted that, generally, the values cherished in scientific cultures (e.g., *humility*, *scepticism* and *conformity*) can be articulated in discoursal meanings by the use of hedges (Hyland, 1997), for example,

DOI:10.6814/NCCU202100169

drawing on the hedge *unclearly* to express scepticism. Specifically, hedge, when looking at the disciplinary community as a whole, tends to be used as a central rhetorical device (a) to achieve the communal acceptance of knowledge claims, given its function of signaling the author's evaluation of the proposition based on standards of assessment in the field of study, and (b) to establish interpersonal solidarity among community members, due to its role basically as a politeness device, which, thus, minimizes the negative interactional effects of information (Aull, 2015; Channell, 1990; Hyland, 1996, 1999; Lancaster, 2016). As to the academic writer as an individual, it is clear that hedges are pragmatically crucial to their expression of uncertainty, scepticism, and demonstration of being open-minded towards one's propositions (Hyland, 1998b), and that by virtue of them, writers can more easily step in the text to instill an attitude towards their claims, readers and disciplinary communities (Chafe & Nichols, 1986; Halliday & Hasan, 2014; Hyland, 1997).

As a result, given such crucial roles that hedges assume in fulfilling the epistemic and interpersonal needs of individual writers and disciplinary communities alike in scholarly writing, it is not surprising that both international and local social scientists display their great willingness to place a hedge between a head noun and its complement, although the main forms of lexico-grammatical realization of hedges are *verbs*, *adjectives*, *adverbs* and *nouns* (see, e.g., Halliday, 1994; Crosthwaite et al., 2017), among which only adverbs are appropriate to be drawn on as post-modification in the *Noun Complement* construction. In other words, writers in the social sciences have plenty of opportunities to place hedges in the form of verbs, adjectives, adverbs and nouns in other positions, such as subject, object, adverbial, or even complement positions, but they still seize every chance to intrude hedges into the text, even in the unusual grammatical slot (post-modification position as in this dissertation), which, thus, reflects the central importance of hedges from another angle. Examples (59a) and (59b) are typical in the two subcorpora:

(59a) Repeatedly disowned, denigrated, and dismissed, it nevertheless refuses to go away — at least circumstantial <u>evidence</u>, **perhaps**, [that there is indeed "some there there."]

DOI:10.6814/NCCU202100169

[SSS INT RA9]

(59b) The assessment of reading got similar answers – **mostly** [that learners did not get any individual marks for reading, except for low-proficiency students, or, alternatively, for students in extracurricular drama classes].

[SSS LCL RA5]

In Example (59a), the hedge *perhaps* is placed in the position between the head noun *evidence* and its *that*-complement clause "that there is indeed 'some there there'". It is apparent here that the writers are taking care to evade the explicit and complete responsibility for the truth value of the *evidence*, for which they not only employ the adverb *perhaps* as the connector but also place another hedge *at least* in adverbial position. In other words, such repeated use of hedges therefore displays the greater reluctance on the part of the writer to demonstrate a strong commitment to the truth of the proposition which is abstracted and packaged into the head noun, and which is then specified in its complement. Similarly, the hedge *mostly* in Example (59b), together with its adjective counterpart *similar* in attributive position, functions to reduce the strength of the proposition nominalized into the head noun *answers* and elaborated in its *that*-complement clause, while also indicating that a pragmatically cautious position is taken by the writer.

Taken together, hedging is universal in scholarly writing, particularly in the soft fields (Fahnestock, 1986; Hinkel, 1997, 2005; Hyland, 2002; Skelton, 1988; Swales, 1998), and its prevalence in research articles from social science journals may in part be attributable to the epistemic and interpersonal needs driven by the shared values of scientific cultures and the idiosyncratic mode of knowledge construction in these specific domains. Hence, social scientists are urged to avail themselves of every opportunity to adopt such a metadiscourse strategy even in the sentential positions where hedges are not routinely expected to emerge. This also explains why local writers in soft domains, although completely disregarding the other two interpersonal forms, still enable the use of hedges in their research articles.

DOI:10.6814/NCCU202100169

The proportion of this pattern of use, however, appears to be smaller in the local subcorpus than the international one, which, in turn, supports the findings in other studies that Chinese writers of English academic articles are found to make more direct, assertive or even authoritative claims than their native counterparts, given their less use of hedging expressions (e.g., Allison, 1995; Feng & Zhou, 2007; Hu et al., 1982; Milton & Hyland, 1999; Xu, 2011). However, concerning academic writing in the Chinese language, its native writers seem to prefer a more cautious and indirect style when stating opinions. This is mainly due to the traditional Confucian beliefs that social harmony and collectivism are more cherished than individualism, or the highly promoted respect for the ideas of the "superiors" (Winfield et al., 2000: p. 332; i.e., those more knowledgeable, wiser and older), which discourages negative assessment and comments (see, e.g., Bloch & Chi, 1995; Chen, 2020; Hinkel, 1994; Shi, 2003). Based on these, it is reasonable to assume that Chinese writers in their English academic writing hold the misconceptions about individualism and respect for others' ideas in Anglo-Saxon norms and practices, or in other words, they are likely to regard such norms and practices in English writing as the exact opposite of that in Chinese writing. Namely, the more they embrace the ideology of collectivism and treat superiors with respect in their writing in the native Chinese language, the less they do so in its English equivalence. In fact, they tend to view collectivism and individualism as polar opposites rather than being on a continuum. As such, they tend to go to the extreme of going on an "unhedging" spree (heavily overusing unhedged claims) as a way to show their professional honesty and integrity, which is, however, based on the misconceived expectations of Anglo-Saxon norms and practices of academic writing. This, therefore, discourages their use of hedges as post-modification in the *Noun Complement* construction.

On the other hand, the use of boosters (e.g., *certainly*, *evidently* and *actually*) allows writers to convey certainty in their claims, thus emphasizing the force of the propositions and demonstrating strong commitment to them. More importantly, boosters aid in instilling confidence and trust in writers into readers by leaving them an impression of authority, assurance and conviction, which, therefore, contributes to the building of solidarity with

DOI:10.6814/NCCU202100169

readers to hold a viewpoint together against other opinions. Additionally, such a metadiscourse device also helps mark writers' involvement with the topic through their emphasized certainty and firm commitment to information (Chafe, 1986; Holmes, 1984; Hyland, 1999b, 2005b).

As such, on the part of international social science writers, they are always attempting to keep the balance between hedges and boosters, granted firstly that they need to coordinate their authority as "expert-knower[s]" (Hyland, 2005b: p. 91) via boostering, and humility as "servants of the discipline" (Myers, 1989: p. 4) through the use of hedges. Secondly, since hedging is employed out of concern for others' negation or face-saving, whereas boosters are for writers' own rhetorical choices to strengthen the certainty or assurance of what they have written, the two are supposed to be carefully balanced in the way that writers' propositions are imbued with convincing degrees of assurance, while also avoiding the overstatement of their claims and thus the risk of the denial by readers. Finally, the combination of hedges and boosters seems like "the iron fist in the velvet glove" (Hyland, 2005b: p. 69), with the writer taking an authoritative persona to express assertiveness by boosters (or show "the iron fist") as part of disciplinary norms of appropriate arguments, and hedging to show due respect to other community members' professionalism in order not to compromise the truthfulness of evaluation ("the velvet glove"). This is proved by the almost equal proportions of hedges and boosters (7.1 vs. 7.8) in the international subcorpus.

Local writers as native users of the Chinese language are, on the other hand, more influenced by Chinese cultural practices, whose roots are firmly planted in Confucian and Taoist traditions (Hinkel, 1994, 1997; Nisbett et al., 2001; Tran, 2006; Tweed & Lehman, 2002), which is, therefore, liable to cause their rhetorical unawareness of the balance between hedges and boosters. Specifically, although certain studies reveal that Chinese learners of English tend to draw on more boosters than expert writers from academic journals in English (e.g., Huang, 2007; Hyland & Milton, 1997; Wang & Jiang, 2019), social scientists in local English journals, compared to most of English learners in Taiwan

DOI:10.6814/NCCU202100169

or mainland China, are normally productive contributors to journals in Chinese as well (for the development and prosperity of the literature in the Chinese language as well as for using published Chinese articles as individuals' qualifications for promotion up the university status ladder), whereby it is reasonable to assume that they steep themselves deeper in Chinese culture, thus adhering more firmly to its traditions, beliefs and practices.

According to Peng and Nisbett (1999), the deeply held sociocognitive belief that "verbal debate and argumentation are not meaningful tools for understanding truth and reality" (p. 15) is highly characteristic of Chinese cultural practices. Put differently, all truth is considered self-evident thus in no need of arguments (Bodde, 1991; Hu & Cao, 2011; Yang, 2013). Instead, knowledge is supposed to be accumulated on the basis of previous convictions and experience (Yang, 2013), and authoritative knowledge (i.e., the resource for creating a communal way of viewing the world) is profoundly respected (Davis-Floyd & Sargent, 1997; Nakayama & Dusenbury, 1984; Tweed & Lehman, 2002), which, in turn, entails "a focus on particular instances and concrete cases" (Nisbett et al., 2001: p. 294) to further justify and support such knowledge. Also, as Matalene (1985) notes, in Chinese writing, the way in which the authors usually "expect the audience to infer meanings rather than have them spelled out is a defining characteristic of Chinese rhetoric" (p. 801). More importantly, in the rhetorical traditions of the Chinese language, there is no need for writers to validate themselves as being knowledgeable, since their capabilities to publish their writing are enough to prove their authoritativeness, credibility and erudition (Hinkel, 1999).

Taken together, these three features of Chinese rhetoric suggest that traditionally it seems unnecessary for Chinese writers to shape an image of expert authority or reinforce the certainty of what they have written through the use of boosters, given their superior status as a published author and the cultural norm of not "spelling out" meanings, needless to say that discovering, persuading others to accept, or denying the truth are absolutely not central to the mode of knowledge construction in the Chinese writing system, thereby further downplaying the role boosters assume in the structure of argumentation built around

DOI:10.6814/NCCU202100169

the truth. In addition, it is not fairly common for writers to place a booster in the position between a head noun and its complement, which, therefore, accounts for local writers' non-use of this metadiscourse strategy as well.

Lastly, attitude marker is, when taking the form of adverb phrases, another interpersonal metadiscourse device that international writers in the soft knowledge fields favor to place in the position between a head noun and its complement, whereas their local counterparts' repertoire of metadiscourse is devoid of this pattern of use. This kind of markers (e.g., *interestingly*, *surprisingly* and *unexpectedly*), as noted above, reflects the writer's affective attitudes toward propositions, such as surprise, acceptance, disappointment, etc., in a way that foregrounds the writer and contributes to his or her adoption of a professionally acceptable persona and establishment of a connection with the disciplinary community (Hyland, 2005b, 2014). This is because the use of these markers is normally based on community-recognized attitudes, values and practices to information. More importantly, this pattern of use is not just to express the writer's attitude to a proposition but rather to persuade the reader through the emphasis upon what readers are supposed to take heed of and how writers prefer them to make a response to it. To illustrate, the following example from the international subcorpus is presented:

(60) Research on the semantics of definites predates the invention of the talking picture, and to date there is a lack of consensus **even** [as to whether *the* carries an existence presupposition (Coppock & Beaver 2012)], let alone agreement concerning uniqueness and familiarity (e.g. Strawson 1950, Birner & Ward 1994, Roberts 2002, Abbott 2008, Elbourne 2013).

[SSS INT RA3]

Here, the use of the adverb *even* as the attitude marker indicates the writer's affective attitude of surprise to the head noun *consensus* with its specification in the following complement, namely, a preposition *as to* with a *whether*-clause in its object position. In this regard, the attitude marker *even* is drawn on to emphasize that readers should attend to the situation that there is no consensus, even the most basic one, over how the definite

DOI:10.6814/NCCU202100169

article *the* is used in the research into the semantics of definites, and that the writer would like them to respond to this also as being unexpected. In so doing, the writer not only conveys his or her position but also "suck[s] readers into a conspiracy of agreement" (Hyland, 2005b: p. 150) in the way that readers might find it hard to argue against what the writer has expressed. Therefore, the use of attitude markers is not simply for the writer to display emotional attitudes, but there is considerable rhetorical effort involved in projecting the writer's persona, establishing a link with the disciplinary community, and persuading the readers, all of which are underlying the surface function of displaying affect.

As such, it is unsurprising that local social scientists completely overlook its use, probably given their attention merely devoted to its superficial role as an affective carrier, which is, however, not compatible with the rhetorical traditions of Chinese academic writing. Specifically, in terms of the rhetoric in Chinese written discourse, the use of affective expressions is usually deemed inappropriate, since firstly it can function to display very individualized and idiosyncratic aspects of a writer's personality, mood and passion, which, thus, conflicts with the collectivist cultural orientation of native Chinese users. Secondly, the serious tone and style of academic written discourse, cherished in Chinese rhetorical traditions, can be undermined when this pattern of use occurs, for such traditions hold that high levels of seriousness in tone and style of writing contribute to the authority of academic discourses, thus crucial to the readers' acceptance of claims.

More fundamentally, classical Chinese rhetoric owes its existence and development to the two cardinal principles in the teachings of Confucius, namely, *Ren* (humanity and love) and *Li* (propriety) (Lu, 2000). Thus, the writer needs to readily acknowledge the presence of readers and treat them equally, fairly and empathetically, which, distinct from "reader-friendliness" in English writing, entails the joint or shared responsibility of both writer and reader for conveying messages in the text. This runs parallel to Hinds's (2001) "reader responsibility", in the sense that readers are even made "primarily responsible for effective communication" (p. 65). Moreover, this rhetorical tradition might relate to the construct of the "low context" culture (Hall, 1976: p. 79) in the field of intercultural communication,

DOI:10.6814/NCCU202100169

where the propositional information is mainly embedded in its context and implicitly delivered, thus placing the primary responsibilities for interpreting the underlying meanings through contextual knowledge on the reader. In other words, it is the reader's job to infer the intentions of the writer by a broad interpretation of the text together with, for example, the attitude, the affective state, the commitment to a claim, etc., based on his or her personal knowledge and individual perception (Gao & Wen, 2019; Kim & Lim, 2013). As such, it makes no sense for Chinese writers to construct an authorial persona that can interact with readers, to establish the interpersonal and social relations between them and other members of disciplinary communities, or to be overtly persuasive and guiding in the development of arguments. Nevertheless, all of these are corresponding to the underlying functions of attitude markers, and thus appear to be of no use for local writers, which, then, further justifies why its employment goes completely unheeded in the local journal articles.

Additionally, Chinese writers tend to maintain a high level of impersonality in the text as a way to establish a more distant relationship with their readers (e.g., Kim & Lim, 2013; Xie & Teo, 2020). In so doing, they in effect attempt to project the self-image as a mysterious, aloof and omniscient knower (extending the identity of the published author as being a self-evidently knowledgeable one in the Chinese rhetorical tradition) who perhaps lives a cloistered life. Therefore, such an enigmatic projection, along with the tactic approval for published writers' erudition, leads readers to form an impression as to how professionally competent and authoritative those authors and their work are.

Taken together, reluctance to give affective expressions, reader responsibility or low context culture, and indifferent and distant self-imaging in Chinese rhetoric of academic written discourse seem to contribute jointly to local writers' complete abandonment of adverb phrases as attitude markers in the position between a head noun and its complement.

5.3.3 Finite dependent clause

Thirdly, international and local social scientists also differ statistically in how they deploy the lexico-grammatical feature *finite dependent clause* in their post-modification use for different functional purposes (see Table 5.10 below). In general, the international

DOI:10.6814/NCCU202100169

	SSS in CIJA	SSS in CCLJA
Abstract entity	30 (28.8)	0 (0.0)
Other human	33 (31.7)	8 (100.0)
Self-sourced	19 (18.3)	0 (0.0)
Self-sourced + Endophoric marker	22 (21.2)	0 (0.0)
Totals	104 (100.0)	8 (100.0)

Table 5.10 Finite dependent clause as post-modification between SSSs in CIJA and CCLJA (% of total)

subcorpus, compared to its local counterpart, includes a significantly higher frequency of finite dependent clauses as post-modification. This could firstly be due to the grammatical complexity of juxtaposing a finite dependent clause (e.g., a relative clause) and a complement clause, as illustrated in Example (61) below:

(61) What has to be made explicit, then, is the tacit assumption **on which much of Kirkman-Brown and Martins' opinion seems to hinge** [that the value of parenthood does not depend on a genetic link between parent and child].

[SSS INT RA30]

In this excerpt, the head noun *assumption* is separated by a relative clause "on which Bruening's argument rests" from its *that*-complement clause "that a Passamaquoddy quantifier must c-command a variable that it binds". As such, this might, according to Biber et al. (1999), place a heavy burden on the reader's short-term memory in the way that they would need to process the long intervening subclausal constituent (e.g., a relative clause) before finally reaching the complement of the head noun. Furthermore, Biber et al. (1999) also note, a large number of head nouns hold the potential to take either a relative clause or a complement clause, but it is rare for those nouns typically occurring with a *that*-complement clause to concurrently take a *that*-relative clause. Additionally, the juxtaposition of a relative clause and a complement clause is, to a certain extent, liable to confuse readers, in particular the non-native English ones, given the two subsumed under the same category of finite dependent clauses with rather parallel syntactic structures.

DOI:10.6814/NCCU202100169

Therefore, the juxtaposition of this kind is not favored especially by local writers, whose reader groups are mostly native Chinese speakers.

Taken together, readers' memory burden, the repulsiveness of relative and complement clauses, and the juxtaposition carrying the risk of confusion have made a concerted effort to persuade local writers in soft fields to disregard the use of finite dependent clauses as post-modification. Nevertheless, a new question occurs to this pattern of use: why do writers from international social science journals still devote a certain proportion of their post-modification use to finite dependent clauses? The explanations are as follows:

First of all, it is more common for a writer to draw on the pre-modification of stance nouns, such as first-person possessives (*my, our*) to aver the stance, as illustrated in the following example:

> Carlin and Hellwig (CH) challenge **our** claim [that voters misattribute responsibility for economic performance in low-savings commodity-exporting (LSCE) countries of Latin America].

On the other hand, Example (62) is typical of the use of finite dependent clauses as post-modification for *self-sourced* stance-making in the international subcorpus:

> (62) Some simple arithmetic using the cost figures in table 1 in connection with the framework sketched allows us to substantiate the claim **we made** [that the key policy question for this literature is not whether police affect crime, but the extent to which police affect violent crime, particularly murder].[11]
>
> [SSS INT RA43]

Now, if the comparison between these two excerpts is drawn, another interesting question is raised: why bother using a relative clause as post-modification rather than a simple possessive as pre-modification? This may firstly be attributed to the exact wording of the stance-taking process. Specifically, the pre-modification *our* to the head noun *claim* in the first example above actually conveys multiple meanings: (a) the claim can be made by one or two of the authors but not all of them, given the multiple authorship in this research article, (b) the claim might not be originally made but rather proved or bolstered

DOI:10.6814/NCCU202100169

by the writers of the present article, thus becoming their claim as well, and (c) all the authors participate actively in the creation of such an original claim. On the other hand, the post-modification *we made* to the head noun *claim* in Example (62) more clearly gets the message across to readers that *we*, as joint writers of this research article, undertake the collaborative effort to create this original claim. Just as importantly, although the post-modification *we made* is syntactically a relative clause, it comprises merely two words, which, thus, avoids imposing a heavy burden on readers' short-term memory, and confusing them with the juxtaposition of a lengthy relative clause and a following complement clause. Example (63) is typical too:

(63) This finding is crucial to an explanation of bipartite negation as a North African areal phenomenon (found in both Arabic and Berber; Lucas 2007) as it makes possible the hypothesis **we develop** [that negation with *ma*…*-š* in Arabic has its origin in the imposition of bipartite negation by native speakers of Coptic on their L2 Arabic after the Arab conquest of Egypt in the seventh century].

[SSS INT RA58]

In this excerpt, the joint authors employ the two-word relative clause *we develop* to make the clear and unambiguous assertion that all of them initiate a coordinated effort to develop the *hypothesis* without placing heavy cognitive burdens on readers.

Furthermore, it is interesting to note that the deployment of finite dependent clauses for *self-sourced* stance-making, given its clausal form, aids writers to partially aver a proposition and thus obscure their responsibility to it through the manipulation of the clausal predicate. To illustrate, the relative clause *that I know* as the post-modification to the *way*, in Example (64), indicates the writer's averral of the stance expressed through the head noun and its elaboration in the following *of* + *ing*-clause. At first glance, it seems to mean that the writer devised such a way.

(64) There is only one way **that I know** [of meeting the argument that romance is reaction], which is the argument of the paper in hand.

[SSS INT RA18]

DOI:10.6814/NCCU202100169

Nevertheless, the writer in this excerpt in fact skillfully manipulates the main verb *know* in clausal predicate position to his advantage, in a way that avoids making an explicit and complete commitment to the truth of the proposition. This is because "the way that I know" conveys a double meaning: on the one hand, "the way that I know" can be "the way that I create", and it may also be "the way that I know others create" on the other. To put it differently, the writer here only takes on an implicit and incomplete commitment to the propositional truth, since he points tactically to a specific "way" that exists in his knowledge base without averring it or attributing it elsewhere. Actually, this comparison between possessive as pre-modification and relative clause as post-modification can be clearer, provided the extract below, where writers assume overt and total responsibility for a proposition, is compared with Example (64) above:

There is only **my** way [of meeting the argument that romance is reaction], which is the argument of the paper in hand.

Example (65) is also typical in the international subcorpus:

(65) This is not a short paper but reading it is the quickest way **I know** [of reducing the risks in writing a profile or a review, a news feature or an opinion column].

[SSS INT RA32]

It can take the form of the possessive as pre-modification as well:

This is not a short paper but reading it is the quickest of all **my** ways [of reducing the risks in writing a profile or a review, a news feature or an opinion column].

Taken together, the use of finite dependent clauses as post-modification allow writers to flexibly claim the ownership of the proposition in the way that they can either provide even the maximum unambiguity of the averral relationship between them and the proposition, or deliberately obscure such a relationship.

Secondly, finite dependent clause, because of its clausal form that offers more grammatical positions (e.g., subject, predicate, object positions) at the clause boundary (cf. phrasal structures, such as prepositional phrase, adverb phrase, etc.), affords writers more

DOI:10.6814/NCCU202100169

opportunities to place grammatical elements with other linguistic and extra-linguistic functions in such positions, as demonstrated in Example (66) below:

(66) This assertion is predicated upon the assumption **which we defend below** [that all firms in the industry have essentially comparable performance outcomes: to win games].

[SSS INT RA58]

In this extract, although the post-modifying relative clause only consists of four words (i.e., *which we defend below*), the writers make full use of each of them in their respective grammatical positions, namely, *we defend* in clausal subject and verb positions for self-sourced stance-making, *below* in adverbial position for endophoric referencing, and *which* as the relativizer for its syntactic role as the relative-clause element. As such, the use of finite dependent clauses as post-modification allows the writer to realize the combination of stance and metadiscoursal functions (e.g., *self-sourced + endophoric marker*) within the boundary of just one clausal structure.

In addition, such a combination leads to the conciseness expected in academic language (e.g., Bacon, 2013; Hyland, 2015; Snow, 2010), given its feature of one grammatical element carrying out two functions, thereby, to a certain extent, challenging Staples and Reppen's (2016) claim that "noun phrase modification allows academic writers to express more complex relationships and abstract ideas more concisely than clausal elaboration" (p. 18). More precisely, the clausal elaboration (e.g., a relative clause) can be as short in length as the noun modifier but serve more functions. As such, local writers' unawareness of the combination of linguistic and extra-linguistic functions served by a finite dependent clause as post-modification and its attendant conciseness is likely to cause their complete disregard of this pattern of use.

Similarly, only writers in international social science journals employ finite dependent clauses to attribute the stance embedded in the head noun and elaborated in its complement to abstract entities, as shown in the following example:

DOI:10.6814/NCCU202100169

(67) Indeed, there appears to be no evidence for the initial assumption **on which Bruening's argument rests** [that a Passamaquoddy quantifier must c-command a variable that it binds].

[SSS INT RA48]

In this excerpt, the writers attribute the *assumption* to an abstract entity, namely *Bruening's argument*. Now, suppose that the writers replace this relative clause "on which Bruening's argument rests" with a pre-modified possessive, thereby turning Example (67) into:

Indeed, there appears to be no evidence for **Bruening's** initial assumption [that a Passamaquoddy quantifier must c-command a variable that it binds].

In fact, seen from the co-text of Example (67) in the research article, it is clear that this initial *assumption* is part of Bruening's theory about the "movement analysis of raising" and thus made by him. Therefore, the comparison between these two extracts brings up the question: why bother to use a long and complex clausal form as post-modification instead of a short and simple possessive as pre-modification to express the same information? In other words, why do the writers put it as "the initial assumption on which Bruening's argument rests" rather than "Bruening's initial assumption", if this assumption belongs to Bruening?

In early literature, Goffman (1967) defines the concept of face as the desire to be approved of (i.e., positive face) and to act without being impeded (i.e., negative face). This is then modified by Brown and Levinson (1987) in their politeness theory to argue that an author is motivated to protect both aspects of their own and readers' face. Based on this theory, the writer would attend to the balance of face needs because interactions are considered as inherently imposing, whereby a large number of Face Threatening Acts (FTAs) are involved. Myers (1989) further points out that speech acts, such as criticism and counterclaim, constitute FTAs in academic writing, which are, then, against not only readers in the same research area but also those in the wider disciplinary community, thus entailing certain linguistic choices by writers to mitigate threats to all these readers' face.

DOI:10.6814/NCCU202100169

Hence, it seems reasonable, in this regard, to assume that the inconvenience of employing such relative clauses of comparatively great length and complicity can be attributed to writers' tactical linguistic choices to alleviate the face threat to readers. Specifically, the use of a relative clause in Example (67) seems like stating that it is not Bruening who lacks evidence to support his initial assumption but it is his argument that does, thus appearing as if the writers transfer the responsibility for the provision of evidence on Bruening to that on his argument. Put differently, it is not Bruening's fault but rather that of his argument for not having any evidence to attest to such an assumption. In so doing, the threat to Bruening's face, as well as that of those who are engaging in this line of research or even from the wider community, who also accept these assumptions that underpin his "movement analysis of raising", is likely to be mitigated. Similarly, Example (68) below is another typical case in the international subcorpus:

(68) An analysis of Cuba's economic realities has served as the cornerstone for the hypothesis **on which the study is based** [that between 1986 and 2009, the growth of Cuba's production sector was subject to dual constraints stemming from both supply- and demand-side factors].

[SSS INT RA3]

In this excerpt, the same question arises again: why the writer took the trouble to employ a relative clause as post-modification instead of a possessive as pre-modification, by which Example (68) is rephrased as :

An analysis of Cuba's economic realities has served as the cornerstone for my hypothesis [that between 1986 and 2009, the growth of Cuba's production sector was subject to dual constraints stemming from both supply- and demand-side factors].

In fact, the entire research article is centering around the test of this hypothesis, which, in turn, entails the writer following a formal procedure of hypothesis testing by (a) stating the research hypothesis as a null (H_0) and alternate (H_a) one, (b) collecting data in a way that correspond to the variables to test the hypothesis (in this case data on Cuba's economic realities), (c) conducting the statistical tests (KMO and Barlett tests), (d) deciding whether

DOI:10.6814/NCCU202100169

H_0 is supported or refuted (refuted), and (e) reporting the findings in the results and discussion sections. As such, it is clear that the hypothesis is proposed by the writer, and then why did he choose not to simply write it down as "my hypothesis" but as "the hypothesis on which the study is based" instead? This, of course, can be first attributed to the writer's strict adherence to the established conventions of impersonality in English academic writing, given the use of this post-modifying relative clause cloaking the authorial presence in the text. More importantly, the disregard of the possessive *my* seems to indicate the writer's reluctance to give a clear and total commitment to making such an assumption, which is, to a certain extent, further proved by no occurrence of any possessive *my* in the whole article. In other words, it looks as if the writer is attempting to convey the information that this assumption belongs to this study once and only once, thus not becoming any part of his long-term epistemic, value and belief systems.

Taken together, the employment of finite dependent clauses as post-modification to attribute stance nouns to abstract entities is not simply for the indication of the information source, given its alternative *possessives*, being simpler and more direct, but for either the mitigation of criticism, counterclaim and the like toward readers, or the avoidance of the explicit and complete commitment to the information expressed by writers. For the same reason, local writers' rhetorical unawareness of the extra-linguistic functions served by finite dependent clauses as post-modification to attribute stance nouns to abstract entities may contribute to their lack of this use.

Last but not least, the employment of finite dependent clauses to attribute stance nouns to other humans is proportionally most common in both international and local subcorpora. On the part of international writers, functionally, it is also not for the purpose of indicating the information source from other human sources but for the rhetorical purposes as that in the case of abstract entities. To illustrate, Example (69) below is typical in the two subcorpora:

(69) We would like to remind readers of the ethical obligation **to which social scientific researchers are expected to adhere** [that they share data with

DOI:10.6814/NCCU202100169

legitimate colleagues who wish to confirm findings that have already been published].

[SSS INT RA30]

It is interesting to note here that the use of other humans, similar to that of abstract entities, also allows the replacement of the relative clause as post-modification by the possessive as pre-modification, which is shown as follows:

We would like to remind readers of **social scientific researchers'** ethical obligation [that they share data with legitimate colleagues who wish to confirm findings that have already been published].

Apparently, the obligation in this extract is imposed on social scientific researchers, and hence placing a third-person possessive in the pre-attributive position of the head noun is the simplest and clearest way to deliver the intended meaning of the writers. Therefore, the use of any other seemingly redundant grammatical elements (e.g., relative clause) can be seen as suggesting the rhetorical effort invested by the writer. Specifically, the stance noun *obligation*, in this case, seems to represent a threat to the negative face (Brown & Levinson, 1987; Goffman, 1967) of the social scientific researchers, most of whom, actually, happen to compose the majority of the readership of the research article in question, since their actions are probable to be unimpeded by writers' reminder of this obligation.

In this sense, the relative clause "of which social scientific researchers are expected to adhere" is thus used to mitigate such a face threat in the ways that (a) the use of the passive voice in the subordinate clause seems to exempt the writers from shouldering the responsibility of imposing such an obligation, for it might not be them but someone else who expect the readers to do so, and (b) the relative-clause predicate comprises the verb "expect" rather than any modal verbs with obligation/necessity meanings (Biber et al., 1999; e.g., *must*, *should*, *need to*, etc.), or in Huddleston and Pullum's (2002) terms, the modal auxiliaries expressing modal necessity. To be exact, those modals are "semantically strong" (Huddleston & Pullum, 2002: p. 177) in the sense that their subjects are normally

DOI:10.6814/NCCU202100169

forced to do whatever needed to be done as dictated in the text, whereas "expect" is a mental verb based on its semantic domain (Biber et al., 1999), used most of the time to denote the mental state of thinking that someone should behave in a particular way or do a particular thing. Therefore, the message it conveys still stays in people's thoughts and is not put into effect and thus unobligatory in reality, whereby personal volition is allowed on the part of the recipients. In other words, the use of "are expected to" rather than semantically strong modal auxiliaries (e.g., *must*) renders the wording of the relative clause less obligatory, which, in turn, seems to grant such readers who are also researchers the power to make their own decisions about whether to fulfil the obligation or not.

Similarly, as shown in Example (70) below, the finite dependent clause *with which many would agree*, on the surface, appears to attribute the stance noun *view* to "many people" (*other human*), but it actually serves other rhetorical purposes.

(70) Furthermore, there is the view **with which many would agree** [that the best way for students to learn to handle economic statistics is by actually trying to do it themselves].

[SSS INT RA63]

Specifically, the writer displays the preference for an extended and structurally complex clausal element as post-modification over a short and simple genitive or adjectival premodifier (e.g., *many people's* or *prevailing*), as illustrated below:

(i) Furthermore, there is **many people's** view [that the best way for students to learn to handle economic statistics is by actually trying to do it themselves].

(ii) Furthermore, there is the **prevailing** view [that the best way for students to learn to handle economic statistics is by actually trying to do it themselves].

Here, the genitive *many people's*, in (i), differs from the relative clause "with which many would agree" mainly in the use of the predicate *would agree* in the latter. The modal verb *would* in the context of Example (70) is defined in the *Oxford English Dictionary* as "used for talking about behaviour that you think is typical", which, thus, translates the clause "with which many would agree" into the one "with which I think it is typical for

DOI:10.6814/NCCU202100169

many to agree". In so doing, the writer presents the proposition in the following *that*-complement clause as an individual opinion instead of a fact, thereby reducing the potential risk of challenging readers' existing assumptions too much. This rhetorical effort is then reinforced by the use of the noun "many" to mean a considerable but indefinite number of people, whereby the writer is able to avoid the criticism from readers for exaggerating the number of the holders of this view to a certain extent.

Similarly, in (ii), the placement of the adjective *prevailing* in the premodifier position of the stance noun *view* thus indicates that the existence of such a widely held *view* is represented as a fact, whereby its assertiveness is not hedged and hence more subject to readers' criticism. Furthermore, the *Oxford English Dictionary* defines the adjective *prevailing* as "most widely occurring or accepted", based on which the use of this adjective premodifier seems not to have the same rhetorical effect exerted by the noun *many* in the relative clause, since the adverb *most*, which denotes superlative degree, renders the meaning of *prevailing* too strong.

On the other hand, writers in the local journals of the social sciences differ markedly from their international counterparts in this pattern of use, as in Example (71):

(71) There is the view **which most patent jurists entertain** [that it will sound less unjust if a technique in an earlier application is used to negate novelty, as it is obviously unfair that a prior inventor, who has not claimed his invention in a patent application, has to get a license from a later applicant for the use of the invention of his own].

[SSS LCL RA18]

Here, the use of "most" rather than "many" and the absence of "would" in the relative-clause predicate are highly likely to expose the reported source of information in the *which*-clause to the risk of challenging readers' assumptions and attracting their criticisms. In other words, local writers, compared to their international counterparts, seem not to draw on finite dependent clauses as post-modification out of rhetorical purposes. It seems reasonable here to assume that such use is perhaps for local writers to demonstrate their

DOI:10.6814/NCCU202100169

higher level of English-language proficiency, because they were usually instructed by their language teachers since their first introduction to English argumentative writing that the more complex the grammar and syntax are, the higher scores they receive in their assessed written work. This phenomenon in English teaching may have its origin in the aesthetic features of Chinese literary writing, which, to a certain extent, encourage writers to use flowery language or big words to demonstrate their literary skills that come to be cherished as markers of intellectual competence (Curran, 2014; Elman, 2000). Some of these writers, according to Du (2020), are even obsessed with the great complexity of grammar and syntax.

Example (72) is also typical in the local subcorpus, which, thus, further proves that this pattern of use by local writers is their demonstration of high English proficiency:

(72) The awareness **that the learner has** [regarding what to do with people and relationships related to LLA] is the focus here.

[SSS LCL RA12]

Here, the relative clause "that the learner has" as the post-modification of the head noun *awareness* can be substituted with the genitive "leaner's" without changing any literal meanings of the words or intended meanings represented by the writer, as can be seen as follows:

The **leaner's** awareness [regarding what to do with people and relationships related to LLA] is the focus here.

In this way, the modifier of the head noun *awareness* does not sperate it from its complement, the *preposition* + *wh*-clause construction, thus simplifying the syntactic structure of the sentence and rendering its meaning easier to be understood.

Taken together, the use of finite dependent clauses as post-modification is not simply for writers to attribute stance nouns to themselves or elsewhere as it appears to be on the surface, but it, in effect, functions to accomplish the complex variety of writers' rhetorical purposes (e.g., to hedge their criticism against other members of the field). However, writers in the local journals of the social sciences show no enthusiasm for its use, perhaps

DOI:10.6814/NCCU202100169

given their unawareness of its underlying functions other than attribution of different kinds. More fundamentally, it may be attributed to their lack of enough familiarity with its value in allowing a writer to bring "an interpersonal element into the meaning" (Halliday & Hasan, 1976: p. 276) and thus maintain the rapport with the readers, most of whom are actually members of the writer's own disciplinary community. Their employment of these subordinate clauses exclusively for the attribution to other humans, ironically, turns out to be the proof of such unawareness, since they are used not out of rhetorical purposes but the demonstration of language proficiency.

5.3.4 Non-finite clause

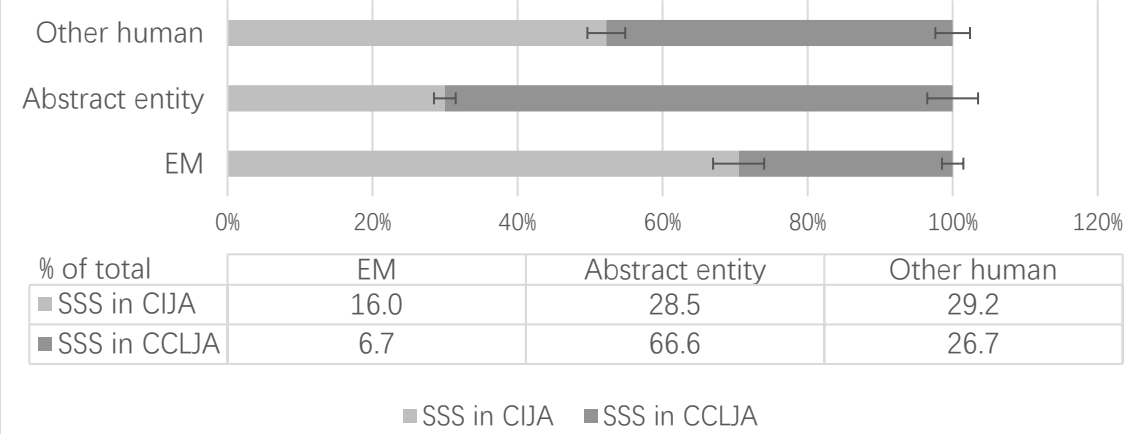
Finally, writers in international and local social science journals demonstrate far different preferences in their use of non-finite clauses as post-modification (see Table 5.11). Proportionally, although there is little divergence in the other human use between international and local writers, as can be seen in Figure 5.5 below, the proportion of abstract entities is nearly two times greater in the local subcorpus than in the international one, and the latter, on the opposite, has almost six times higher density of endophoric markers than the former.

As such, these two groups of writers, proportionately, differ most significantly in their use of non-finite clauses as post-modification to refer to other parts of the text. To illustrate, the non-finite clause *made in this Part*, in Example (73) below, refers the *argument* to "this Part" of the research article. It is interesting to note here that it appears to be grammatically simpler and even more succinct for the writer to use a prepositional phase (e.g., *in this Part*) as the post-modification of the head noun *argument* than the non-finite clause "made in

	SSS in CIJA	SSS in CCLJA
Endophoric markers	54 (39.4)	2 (6.7)
Abstract entity	54 (39.4)	20 (66.6)
Other human	29 (21.2)	8 (26.7)
Totals	137 (100.0)	30 (100.0)

Table 5.11 Non-finite clauses as post-modification between SSSs in CIJA and CCLJA (% of total)

DOI:10.6814/NCCU202100169



% of total	EM	Abstract entity	Other human
SSS in CIJA	16.0	28.5	29.2
SSS in CCLJA	6.7	66.6	26.7

Note: EM = Endophoric markers.

Figure 5.5 Non-finite clauses as post-modification between SSSs in CIJA and CCLJA

this Part", although the latter is just one word longer than the former. However, given the Anglo-Saxon norms of academic writing, reflected in part in the conciseness of the written work, the length of a clause in English is not supposed to be longer than necessary for readers to understand what writers intend (see, for example, Ariana, 2010; Bizzell, 1992; Demir, 2019; Du, 2020), which, strictly speaking, even renders the redundancy of one word unacceptable. Therefore, a question arises here as to why the writer made such a one-word redundancy through a non-finite clause, instead of attempting to avoid it by deploying a prepositional phrase.

(73) The reliance on neighbors' recollections and similar sources reinforces the argument **made in this Part** [that other legal and social institutions supplemented metes and bounds descriptions, making them less inscrutable to the inhabitant of colonial New Haven than they appear today].

[SSS INT RA34]

This might be due to the syntactic structure of the non-finite clause *made in this Part*, in the sense that the agent of the past participle *made*, though omitted as it often happens in passive constructions (the non-finite clause in question, as Biber et al. (1999) note, can be paraphrased as "which is made in this Part", thus a passive construction), is often

DOI:10.6814/NCCU202100169

assumed to be the writer. More importantly, the omission of the agent in this case can be attributed to the fact that there is no need for its identity to be stated, since readers will always infer that the writer is the agent of the action if no identification of another agent (e.g., by other jurists) is made. As such, this means that the writer attributes such a previously made "argument" to himself again, which, in turn, seems to reaffirm his clear and complete commitment to the truth of this argument, thereby instilling the writer's considerable degree of confidence in it once more. From the perspective of readers, this offers them a more convincing degree of assurance about this argument, or more broadly, leaves them an impression as to the certainty, effectiveness and credibility of the writer's claim. In this sense, the addition of just one word *made* cannot be understood as simply turning a prepositional phrase into a non-finite clause from a grammatical point of view, but rather reflects all the rhetorical effort that the writers has invested to convey what he intends.

Similarly, in Example (74) below, the use of the prepositional phrase *in the prior section* can perfectly refer readers to the *evidence* presented earlier in the article, thus appearing to render the past participle *presented* redundant. However, the incorporation of *presented*, likewise, is not out of the pure grammatical purpose but serves the same

(74) In this sense, the instrument builds on the evidence **presented in the prior section** [that banks reallocate credit supply to non-tax-changing states in response to tax shocks occurring in other states].

[SSS INT RA38]

rhetorical function of satisfying readers of the truthfulness of what the writer conveys as in the case of *made* above.

(75) In this light, recall the Restatement provision **quoted earlier** [that a "trustee's decisions ordinarily must not be motivated by a purpose of advancing or expressing the trustee's personal views concerning social or political issues or causes]."[174]

[SSS INT RA4]

DOI:10.6814/NCCU202100169

Furthermore, the employment of non-finite clauses as post-modification, other than for the rhetorical purposes mentioned above, is to ensure the accuracy of the conveyance of information about the reference to other parts of the text. To illustrate, suppose, in Example (75), that the past participle *quoted* is omitted from the non-finite clause, which, then, transforms the extract into:

> In this light, recall the Restatement provision **earlier** [that a "trustee's decisions ordinarily must not be motivated by a purpose of advancing or expressing the trustee's personal views concerning social or political issues or causes]."[174]

Here, the omission of *quoted*, to a certain extent, reduces the accuracy of the conveyance of information regarding the post-modification of the stance noun *provision*, granted that whether the provision is made or quoted (actually it can be done in any possible ways) by the writer seems to be unclear, although the use of just "earlier" can perfectly refer readers to the previous text. More importantly, such conveyance of vague meaning might influence the efficient processing of the text by readers, since they are, given the first occurrence of this "provision" being four pages earlier in this article, likely to retain a rather dim memory of its source (from the writer or others), which, in turn, means that they may have to turn back to this previous page and check for the source of information. Example (76) is also typical in the international subcorpus:

> (76) Next comes the modified diachronic path **repeated below in (46)** [in terms of how we analyze VoiceP for each *se* construction].
>
> [SSS INT RA32]

In this excerpt, it seems clear that the exclusion of *repeated* from the non-finite clause leaves the function of referring to other parts of the article through the post-modification intact, but causes the conveyance of meaning to be unclear:

> Next comes the modified diachronic path **below in (46)** [in terms of how we analyze VoiceP for each *se* construction].

Specifically, the construction "path below in (46)" here is highly likely to mislead readers to consider the "path" as a newly forged one, simply because it is located "below

DOI:10.6814/NCCU202100169

in (46)", or in other words, the adverb *below* is used to cataphorically predict that the "path" is going to be specified (usually for the first time) later, with *in (46)* revealing its exact future location.

Taken together, local writers' rhetorical unawareness of the underlying functions performed by non-finite clauses as endophoric markers in post-modification structures is probably held responsible for lowering the proportion of this pattern of use in the local subcorpus.

On the contrary, there are over two times as many non-finite clauses used for the attribution to abstract entities in the local social science subcorpus, as compared with its international counterpart (see Figure 5.5 above). This is mainly due to the rhetorical traditions of the appropriate expression of criticism in the Chinese language. Specifically, according to Bloch and Chi (1995), there is almost no difference in the frequencies of criticizing other studies between Chinese writers and their Anglo-American counterparts in the social sciences. Or, in their own terms, a similar "critical edge" (p. 249) also exists in the texts written by Chinese scholars, given that the emphasis on *thinking* rather than merely *memorizing* in the teachings of Confucius encourages critical reasoning as well. However, Chinese writers' preference for rhetorical strategies of indirectness help evolve their more cautious and implicit style in the process of critically assessing the value of others' work (Kaplan, 1972; Liu & Du, 2018; Scollon, 2001; Scollon & Scollon, 1995). Given this, attributing the source of the stance noun to abstract entities (e.g., the data or the results) instead of referring it directly to other humans aids them in concealing the source of the evaluation. It seems as though the writers are not undertaking the critical assessment of a real life human with emotions, perspective and beliefs, but rather an abstracted and neutral subject. In so doing, the threats to the face of those criticized are mitigated, while also demonstrating writers' goodwill gesture to maintain the rapport with them. Example (77) is typical in the local social science subcorpus:

(77) This finding is inconsistent with the <u>statement</u> **issued by the White House Office of International Economic Affairs** [that China-U.S. trade hurts U.S. interests].

DOI:10.6814/NCCU202100169

[SSS LCL RA17]

Here, the post-modifying non-finite clause "issued by the White House Counsel for International Economic Affairs" attributes the stance noun *statement* to an abstract entity, namely, *the White House Office of International Economic Affairs*. In fact, the content of the *statement* specified in the following *that*-complement clause (i.e., that China-U.S. trade hurts U.S. interests) is a direct quotation from a late American economist *Paul Samuelson* (i.e., *Samuelson's Concern*), which was, then, spoken by the then office deputy director *Kenneth Juster* in his statement at the White House news conference on Monday January 23, 2017. However, neither of these two names are incorporated into the writer's work as cited authors. Actually, there is even no occurrence of any of these names in the entire article. This excerpt may have been simplified as:

> This finding is inconsistent with the statement **issued by Samuelson** [that China-U.S. trade hurts U.S. interests].
>
> Or:
>
> This finding is inconsistent with the statement **issued by Juster** [that China-U.S. trade hurts U.S. interests].

As such, the writer's choice of the attribution of sources to abstract entities rather than to other humans suggests how he maintains rapport and mitigates criticism.

Moreover, Example (78) is another typical case in the local subcorpus:

> (78) It is the difficulty in supervising and judging the safety of genetically modified foods effectively and the uncertainty of the risks **reported in the literature of the food industries** [that GM foods can cause the development of diseases which are immune to antibiotics].

[SSS LCL RA40]

In this extract, the writer attributes the stance noun *risks* to an abstract entity *literature* through the non-finite clause "reported in the literature of the food industries". It is interesting to note in this case that the text writer does not directly conduct the critical assessment of GM foods, whereby she should have found it unnecessary to satisfy the face

DOI:10.6814/NCCU202100169

needs of others who advocate these foods. Actually, the criticism of GM foods being risky comes from the sources in the literature of the food industries. The writer, however, does not cite any authors who hold the opinion as to the risks of GM foods from such literature here. Therefore, it is reasonable to assume that the writer, though not as the direct sources of the criticism, still prefers the attribution to abstract entities over other humans, granted that the absence of cited authors enables her to obscure the origin of negative evaluation performed. This makes it more difficult for readers to know exactly which studies the writer refers to and to challenge them if necessary, especially when the results from different research are mixed and thus inconclusive, as in this case where the safety of GM foods is still under heated debates, with diverse interests involved. To conclude, the attributions made to abstract entities instead of other humans, when the debate topics are highly controversial or sensitive, also help avert the counter-attacks from the readership against the cited authors, which can, then, be seen as the extension of Chinese writers' cautious and implicit style evolved from their rhetorical strategies of indirectness.

Lastly, international and local social scientists make similar use of non-finite clauses for the attribution to other humans. This is not surprising since writers from the soft knowledge domains tend to build arguments and knowledge based on their personal interpretations and negotiations with readers, whereby new arguments and knowledge are not advanced solely from writers' own thoughts and ideas but rather "follow[] altogether more reiterative and recursive routes as writers retrace others' steps and revisit previously explored features of a broad landscape" (Hyland, 2004: p. 31). As such, given the vast literature open to greater interpretation, the key findings more heavily borrowed from related fields, and the inconsistent and unclear criteria for supporting and denying claims, readers of social science journals are usually not assumed to have the same interpretative knowledge, whereby writers need to establish and elaborate a context through citations (Hyland, 2004, 2005b). The more frequent citations in the text, according to Hyland (2014), demonstrate writers' stronger commitment in "firmly situating research within disciplinary understandings, providing a discursive framework for arguments and demonstrating a

DOI:10.6814/NCCU202100169

plausible basis for claims" (p. 31). In this respect, the use of non-finite clauses for attributing to other humans is ideally to suit this very need of writers in soft fields, since it often takes the form of integral citations, a practice of inserting the names of cited authors into the body of the sentence (Gray, 2015; Maher, 2015; Thompson, 2001, 2012), which, in turn, affords them another linguistic option to incorporate as many cited authors into their own work as possible.

More importantly, this pattern of use also allows writers to attribute ideas and insights to prominent disciplinary figures (e.g., *Tony J. Silva* in the field of L2 writing), thus bringing themselves into an alignment with a particular scholarly community. In so doing, their new interpretations and evaluations, given having been embedded in the literature of the community, can then exhibit their "relevance, importance and the credentials of the writer" (Hyland, 2005b: p. 158). Examples (79) and (80) are typical in the two subcorpora respectively:

(79) Indeed, in Robinson's (2002:76) corpus study of Tenejapa Tseltal, he states that he has 'found little evidence in favor of the claim **first made by Smith (1975) and later cited by Norman and Campbell (1978) and Dayley (1981)** [that Tenejapa [Ts]eltal constituent order is determined by a hierarchy of animacy]'.

[SSS INT RA27]

(80) Some believe that the right reduction of the limitation of action system deviates from the general moral sentiments and natural justice principle of human beings, and the opinions **held by some scholars** [that the limitation of action safeguards public welfare, maintains social order, guarantee trust protection and promotes social efficiency and that it does not protect those sleeping on rights cannot justify the limitation of action].

[SSS LCL RA12]

Additionally, Chinese writers are, as noted earlier, obsessed with lengthy and complex sentences, and grammatical features such as clauses, among which the use of non-finite clauses are usually viewed as superior to that of finite clauses. This originated from the

DOI:10.6814/NCCU202100169

Feature	Mann-Whitney			Effect size (*r*)
	U	Z	*p*	
Punctuation	**3879.0**	**-9.971**	**< .001**	***r* = -0.576**
Adverb Phrase	**6817.5**	**-6.299**	**< .001**	***r* = -0.357**
Finite dependent clause	**8093.5**	**-5.581**	**< .001**	***r* = -0.322**

Table 5.12 Lexico-grammatical features of significant inferential statistical differences (in descending order of effect size *r* value) between NSS and SSS in CIJA

emphasis on the use of "fancy syntax" (Enfield, 2009: p. 39) and big words as a way to show one's literate skills in the English classes they had taken before in school. As such, granted that non-finite clauses are compressed in structure and less explicit in meaning, which, in turn, requires higher English language proficiency to correctly compose them, they are more favored by Chinese writers to showcase their literate skills. In this sense, although Chinese writers tend to display reluctance to make attributions to other humans, they are, given the demonstration of high language proficiency embedded in the use of non-finite clauses, more encouraged to employ such attributions.

5.4 Explanations for different use between international writers in the natural and social sciences

The use of punctuation, adverb phrases, and finite dependent clauses as post-modification demonstrates significant statistical differences between international natural and social sciences writers (in descending order of effect size *r* value; see Table 5.12). As such, the remainder of this section is going to unfold first with the analyses of punctuation.

5.4.1 Punctuation

Table 5.13 and Figure 5.6 below compare the use of punctuation as post-modification in the international subcorpora of natural and social sciences. Proportionately, the widest divergence of this use between the two subcorpora lies in the deployment of semicolons, with that of dashes, colons and commas becoming progressively less divergent.

Semicolon, as discussed earlier, can function as a break between two closely related sentence parts, usually independent finite clauses, whereby the clause following the

DOI:10.6814/NCCU202100169

	NSS in CIJA	SSS in CIJA
Comma (,)	42 (40.0)	379 (37.8)
Colon (:)	49 (46.7)	342 (34.1)
Dash (—)	14 (13.3)	269 (26.8)
Semi-colon (;)	0 (0.0)	13 (1.3)
Totals	105 (100.0)	1003 (100.0)

Table 5.13 Types of punctuation used between NSS and SSS in CIJA (% of total)

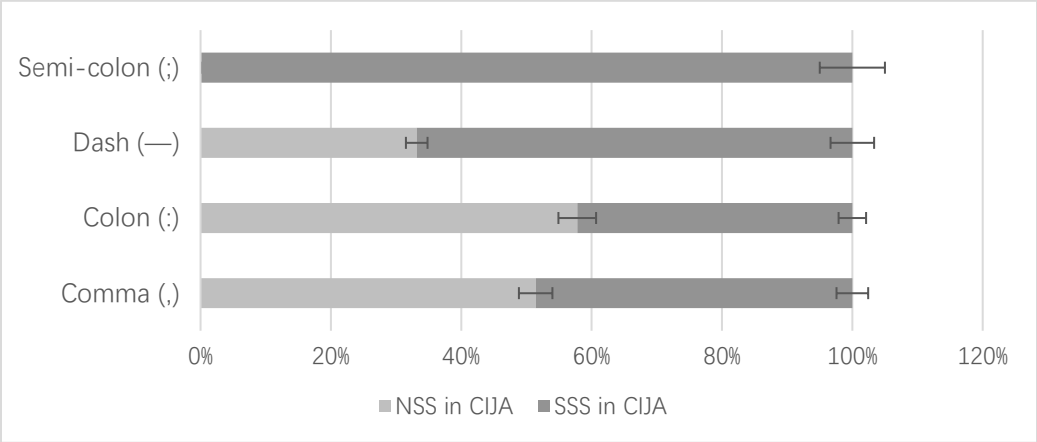


Figure 5.6 Punctuation as post-modification between NSS and SSS in CIJA

semicolon is able to elaborately interpret the one that precedes it. Although semicolon, when used for elaborative interpretations, is stylistically more formal than comma and dash (both of which can perform the same interpreting function), its less use as post-modification on the part of international natural scientists is not supposed to be seen as indicating the lower levels of formality in their work. In fact, the syntactic complexity and the structural type of the clause that immediately follows the semicolon mainly determine the writer's deployment of it in their post-modification use.

Specifically, the incorporation of semicolons into the text entails its following clause to be both an independent finite clause and of considerable length and complexity. However, in the hard knowledge domains, there is, according to Biber (2006a), a growing tendency towards phrasal complexity wherein clauses are maximally compressed into complex phrases (e.g., pre-modifying a noun by a different noun, instead of post-modifying

DOI:10.6814/NCCU202100169

it with a relative clause). In contrast, social scientists favor the extensive use of the dependent clause, or in Biber et al.'s (2011) terms the *T-unit*, namely, "a main clause and all associated dependent clauses" (p. 7).

Actually, following Biber and his associates' work, a scholarly consensus has been achieved that the overall sentence structure is lengthier and more complex in the soft than hard sciences. This can be attributed to its emphasis on the incorporation of multiple dependent clauses, normally with multiple levels of embedding, as a way to render writers' argumentation elaborated and explicit, thereby fitting the mode of knowledge construction based on plausible reasoning and explicitly interpretative persuasion in such domains (e.g., Biber & Gray, 2013, 2016; Egbert, 2015; Gray, 2015; Staples et al., 2016). On the other hand, knowledge in the natural sciences is typically constructed on hard facts and empirical evidence through experimentation and scientific observations, which, in turn, entails a more linear and production-oriented approach, therefore "allow[ing] arguments to be formulated in highly standardised, almost shorthand, ways" (Hyland, 2008: p. 19).

In this sense, it is not surprising that the phrasal complexity which pertains to the maximal compression of clauses into complex phrases, despite its associated loss in elaboration and explicitness of meaning, prevails in the hard knowledge fields. The argument structure involved in such fields is linear and standardized, or even in a shorthand way, whereby readers (most of whom are also specialists in the same field) are usually assumed to be familiar with such routine knowledge and practices by writers, who, of course, choose not to bother themselves with elaborated sentence structures and clausal complexity. Not to mention, in so doing, they actually conform to one of the Anglo-Saxon norms of scientific writing, namely, *conciseness*. As a result, the occurrence of long and complex sentences, usually with multiple levels of clausal embedding, seems to be rather uncommon in hard science research articles, which, thus, contributes to the extremely limited use of semicolons by natural scientists in the international journals, due to the requirement for its following clauses to be of great length and complexity.

DOI:10.6814/NCCU202100169

More importantly, the clause placed right after a semicolon is commonly employed to elaborately interpret the clause immediately preceding it, with the two demonstrating a closely-related explanatory relationship. Hence, natural science writers may, in view of the relatively low demand for elaboration and explicitness of meaning in their fields, find it unnecessary to go through an overly complicated process of placing an adequate explanation or interpretation in the form of a clause in the position immediately after a semicolon. In particular, this is done to further elaborate another *clause* just prior to this semicolon or enable it to be more explicit. All these, then, result in even less deployment of semicolons in the hard field articles.

Additionally, another grammatical constraint of the semicolon use is that in general, only *independent* finite clauses are qualified to be placed in the position right behind the semicolon. As such, even in the international journals of the social sciences, writers are reluctant to allow a semicolon to connect the stance noun and its complement (e.g., *that*-clauses, *of* + *ing*-clauses, *or* wh-clauses), simply because all such clauses are finite or infinite *dependent* clauses, unqualified as the grammatical elements for elaborative interpretations through semicolons, needless to say their counterparts in the natural sciences.

Concerning the use of dashes, writers from international soft science journals demonstrate a clearer preference for it than their hard field equivalence. Dash, as noted above, assumes the role in introducing the grammatical elements that immediately follow it to further elaborate or explain what right precedes it. In this function, dash often competes with comma, whose advantage lies in its higher level of formality, which, then, suits it better to the genre of research articles. However, although construed as "the least formal" punctuation mark, dash has its merits in both marking an evidently stronger break from the surrounding text and allowing a richer variety of clausal structures to follow it than comma.

Specifically, in terms of formality, van Maanen (1995), in early literature, identified a general long-term trend towards greater informality in academic writing, where such a

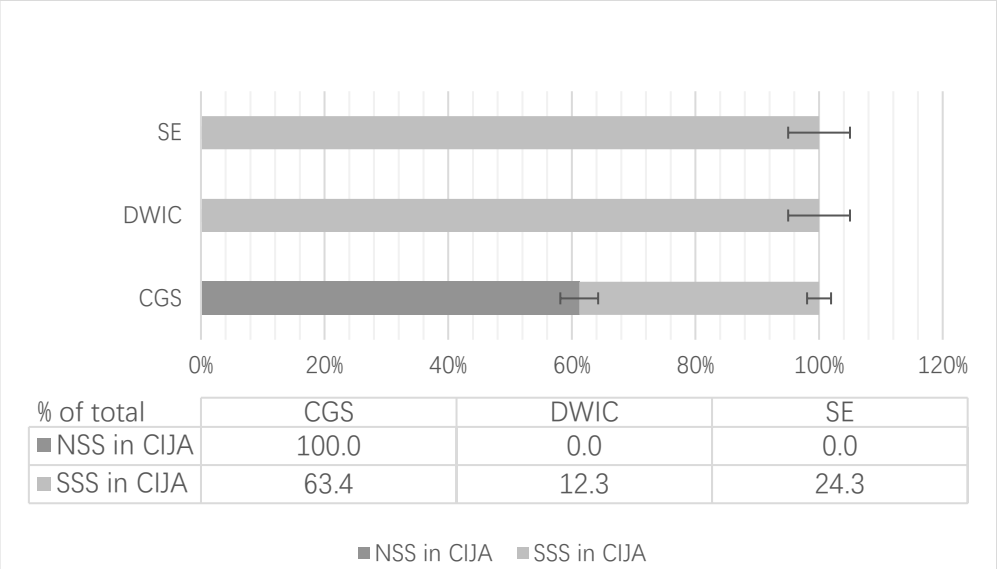
DOI:10.6814/NCCU202100169

shift is “certainly abetted by post-modern and feminist movements, in which the scholarly voice is deliberately disassociated from detachment and distance” (as cited in Swales et al., 1998: p. 118). In this regard, it is apparent that post-modern and feminist movements, in general, have been directly attached to the soft fields of *philosophy*, *sociology*, *literary criticism*, etc. (see, e.g., Ashenden, 1997; Garry & Pearsall, 2015; Vidal, 1991), which, therefore, seems to exert a more profound influence on social scientists, thus leading to the increasing levels of informality in their work. More recently, Cameron (2012) attributes this general informalization trend in scholarly writing to the factors that are “fundamentally social and indeed political” (p. 252), in a way that further distances this movement from the hard knowledge domains. Actually, she also notes that “impersonality still rules in the ‘hard’ sciences” (p. 252), which, given the close correlation of impersonality with informality (see, e.g., Sigley, 1997), proves the more formal and impersonalized style in natural science research articles.

Taken together, the shift to informalization driven by post-modern and feminist movements as well as the social and political reasons that push social scientists more into the informal and personalized writing style on the one hand, along with the ruling of impersonality or formality still in the hard fields on the other, contributes jointly to the discrepancy in formality between these two knowledge domains. This, thus, determines the merely halved use of dashes proportionally by natural scientists, compared to their social counterparts.

On the contrary, writers who contribute to the international journals of the natural sciences tend to make more use of colons as post-modification than their counterparts in the soft fields (see Figure 5.7 below). Nevertheless, all of their colon use is devoted to introducing complex grammatical structures (CGS) that immediately follow this punctuation mark. Actually, colon, as noted above, usually occupies the role as the indicator in forecasting the upcoming coordinated clauses, with two exceptions where (a) the meaning of the head noun is of central importance to the writer’s full delivery of all

DOI:10.6814/NCCU202100169



% of total	CGS	DWIC	SE
NSS in CIJA	100.0	0.0	0.0
SSS in CIJA	63.4	12.3	24.3

Note: CGS = Complex grammatical structures, DWIC = Dependent *wh*-interrogative clauses, SE = Substantive expressions

Figure 5.7 Types of colon use between NSS and SSS in CIJA

core messages in the text (SE), and (b) the following grammatical elements happen to be dependent *wh*-interrogative clauses (DWIC).

Firstly, in terms of CGS, one may wonder why the prevalence of phrasal rather than clausal complexity in hard fields is able to result in the more frequent occurrences of such clausal structures. In effect, the CGS in question refer to the grammatical elements consisting of at least two coordinated clauses. Given that the mode of knowledge construction in the natural sciences, as discussed earlier, relies on hard facts and empirical evidence collected through experimentation and scientific observations, science production is the result of a steady linear and atomistic progression where by "linear", it refers to the development of arguments in a highly standardized way, while "atomistic" entails separating the complex whole into discrete pieces for analysis (Becher, 1994; Becher & Trowler, 2001). In this sense, the way that the analysis in hard fields is usually based on the parts of a whole and also conducted in a systematically patterned method seems to render the use of coordinated clauses necessary. Specifically, each coordinated

DOI:10.6814/NCCU202100169

clause can be employed to correspond to one discrete piece of the complex whole, every time the writer starts to explain a new "whole" in the research article. Since a clausal structure offers a broader variety of grammatical positions (e.g., indirect object, adjunct, qualifier, etc.) to incorporate more meaning elements than the phrasal one, it is able to draw natural scientists' attention especially when they find it difficult to fully elaborate or explain the discrete piece through a phrasal construction (e.g., a noun phrase with multiple modifiers).

Moreover, as the English language allows ellipsis in coordinated clauses that share grammatical elements with a preceding clause, the deployment of such clauses seems not to hinder natural science authors from maintaining the conciseness of argument, which is, as noted above, deemed to be one of the Anglo-Saxon norms of scientific writing. Rather, this pattern of use aids them in meeting the requirement of expressiveness in the conveyance of information about abstract and complicated entities, concepts or processes in their hard science writing. Last but not least, Lim (2017) notes the particular function of coordinated clauses in the description of experimental procedures in laboratory research reports. Namely, the clauses of this type can be drawn on "to display a cautiously constructed sequence of steps taken, thus signalling a higher level of precision in descriptions" (p. 68). In this sense, the use of coordinated clauses in the experimental research reports of the hard sciences contributes greatly to the clearer and more orderly sequencing of methodological steps, which, in turn, leaves readers with the impression as to the writer's highly precise description of experimental procedures. Given the greater popularity and prevalence of experimental research reports in the hard knowledge domains, it is reasonable to assume that natural scientists are more liable to place coordinated clauses in the complement position separated by a colon from its head noun.

As to the use of SE, writers in hard fields, given their far stronger preferences for phrasal complexity and sole enthusiasm for coordinated clauses in terms of clausal complexity, are much less likely to employ other clausal elements to specify or explain stance nouns, even if the substantiveness of these nouns appear to be particularly noticeable.

DOI:10.6814/NCCU202100169

Not to mention, this pattern of use entails the clausal structure placing after the colon as post-modification of a stance noun. Instead, it seems to be more probable for them to draw on a simple sentence characteristic of phrasal complexity for the elaboration or interpretation of the substantive stance noun in question.

In the cases of DWIC, Biber et al. (2002) identified a dimension of variation across different spoken and written registers, with a group of linguistic features associated with the dimension and assigned positive and negative scores. In this regard, the higher the positive loading, the closer a text is to the "affective, interactional, and generalized" end of the dimension, while the lower the score, the closer to its "high informational density and precise informational" end (p. 17). The results show that *wh*-clause is one of the most distinctive linguistic features with positive loadings. Following this, Weigle and Friginal (2015) further point out that "papers in the natural sciences and engineering tend[] to load on the negative pole [of the dimension]", or to put it another way, natural scientists are much liable to render their work most informationally dense and precise. This, to a large extent, rules out the possibility of using positive-loading-assigned *wh*-clauses (to which dependent *wh*-interrogative clause belongs) in their work, which, in turn, explains their non-use of dependent *wh*-interrogative clauses after colons as post-modification.

Finally, the proportion of the comma use in the international subcorpus of the social sciences is parallel to that in its natural sciences counterpart. Notwithstanding this, it is clear from Figure 5.8 below that functionally, the use of commas by international natural scientists is exclusively for introducing the identity relationship, while although this pattern of use also dominates in the contrasting social sciences subcorpus, almost one fourth of the comma use here is devoted to writers' launch of strategic effort.

As noted earlier, given the increasingly lengthy and complex content of apposition, there is a tendency for academic writers to place the nominal complement clause (which, to a large extent, provides more grammatical positions to include meaning elements than phrasal structures) in apposition position as an alternative to the appositive noun phrase. Moreover, the apposition is commonly introduced to its head noun by a comma as a way

DOI:10.6814/NCCU202100169

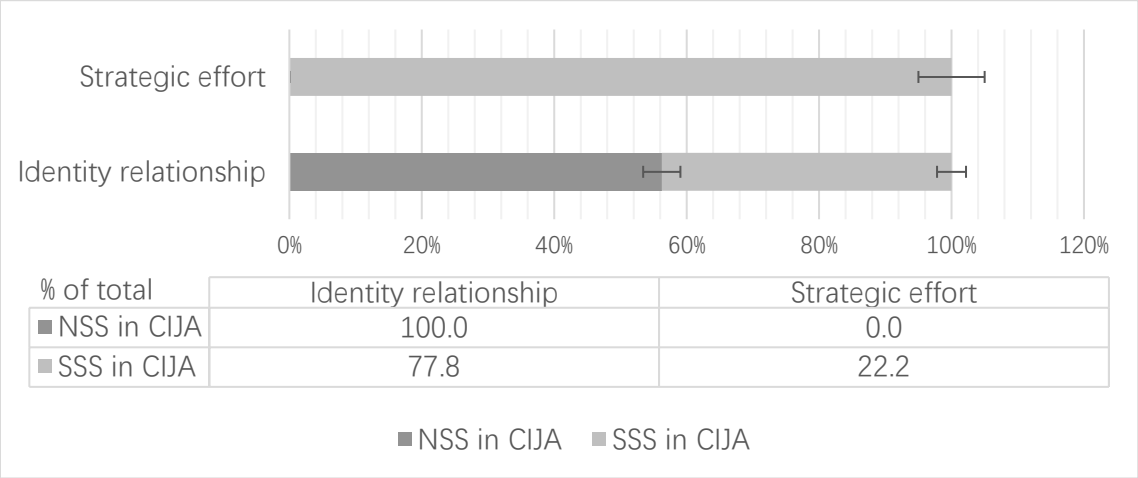


Figure 5.8 Types of comma use between NSS and SSS in CIJA

to show the identity meaning relationship between head noun and apposition, or in other words, it is indicated by a comma that its following apposition is about to clarify the identity of or add descriptive information about the head noun. According to Biber and Gray (2016), appositive noun phrases as noun post-modifiers "have increased in use considerably over the past one hundred years" and are now "extremely frequent in academic texts" (p. 203), while they "can be quite long and complex" (p. 205) as well. Hence, it becomes far more likely than ever before that academic writers need to turn to the nominal complement clauses when the noun phrases cannot provide adequate grammatical positions to include all the necessary meaning elements for the complete conveyance of information from them. As the complement clause competes with and increasingly replaces the noun phrase to demonstrate such a meaning relationship in apposition position introduced by a comma, there is no wonder that the comma use has become dominant in the international subcorpora of both natural and social sciences.

Furthermore, the differences in the growth of knowledge that characterize the soft and hard fields explain the non-use of commas for the launch of writers' strategic effort by natural scientists. Specifically, in the soft knowledge domains, the mode of knowledge

DOI:10.6814/NCCU202100169

construction is, as discussed earlier, grounded in plausible reasoning and explicitly interpretative persuasion, which, in turn, entails a co-production of the work by the writer and by the members of the audience to which it is directed. This is because either for the reasoning to be accepted or the persuasion to succeed, writers are supposed to strategically bring readers into alignment with their points of view and establish a common ground between the two parties. The employment of commas for the launch of writers' strategic effort, as noted above, refers mainly to writers' strategic manipulation of this punctuation mark to entice readers into their stances, which, then, contributes to the establishment and development of their argumentation. On the other hand, knowledge in hard fields is largely constructed on empirical evidence and origination of facts through experimentation and scientific observation. In this way, natural scientists have developed a less reader-inclusive relationship between writer and reader in their work which, then, places greater stress on "the impartiality and linearity of science production" (Hyland & Jiang, 2016: p. 32). As such, natural science writers in international journals are far less liable to draw on commas for their strategic effort to coax readers into their viewpoints as the way to establish and develop their argumentation, thus explaining why proportionately they devote zero percent of their comma use for such effort.

5.4.2 Adverb phrase

Table 5.14 and Figure 5.9 below compare the use of adverb phrases as post-modification between writers from international journals of the hard and soft sciences. Overall, while proportionally, the international subcorpus of the natural sciences contains a significantly higher percent of adverb phrases used for both code glosses and endophoric markers, none of its adverb phrases is devoted to the execution of any other metadiscourse strategies, as compared to its contrasting social scientific subcorpus, where writers employ adverb phrases as both interactive and interpersonal metadiscourse markers. Specifically, the most common function of code glosses is, as previously stated, often used to signal or provide cues to the extra information in the complement position of the head noun for readers, grounded in how writers evaluate the existing knowledge base of their readership.

DOI:10.6814/NCCU202100169

	NSS in CIJA	SSS in CIJA
Textual (*Interactive*)	*56 (100.0)*	*308 (77.5)*
Transition markers	0 (0.0)	63 (15.8)
Frame markers	0 (0.0)	39 (9.8)
Endophoric markers	9 (16.1)	2 (0.5)
Code glosses	47 (83.9)	204 (51.4)
Interpersonal	*0 (0.0)*	*89 (22.5)*
Hedges	0 (0.0)	28 (7.1)
Boosters	0 (0.0)	31 (7.8)
Attitude markers	0 (0.0)	30 (7.6)
Totals	56 (100.0)	397 (100.0)

Table 5.14 Adverb phrase as post-modification between NSS and SSS in CIJA (% of total)

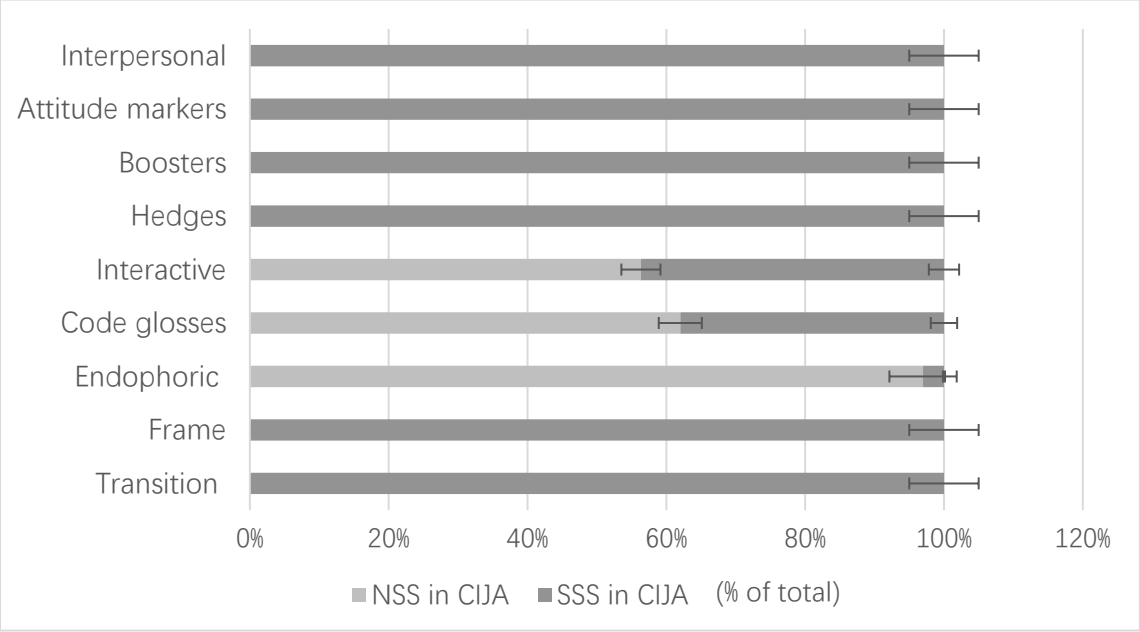


Figure 5.9 Adverb phrase as post-modification between NSS and SSS in CIJA

Besides, the use of such a metadiscourse marker allows writers to make rhetorical effort either to establish authority or enable readers to be aware of and retain the key messages that their work disseminates.

In a review study of how metadiscourse has changed in scholarly writing through the past 50 years, Hyland and Jiang (2018) note a considerable growth in the deployment of

DOI:10.6814/NCCU202100169

code glosses since 1965 by both natural and social scientists. Specifically, although this metadiscourse device has risen by 35% in its frequencies in the soft knowledge domains, it has experienced a far more significant increase of 74% in the hard fields. This can be first attributed to the increasing complexity of scientific research (see, e.g., Ernst, 2005; Fanelli, 2020; Kwiek, 2012; Akin & Scheufele, 2017). To illustrate, the globally competitive environment for scientific and technological innovation entails the greater use of code glosses than before in shedding light on or paraphrasing the cutting-edge concepts, ideas, or terms considered by writers not included in the existing knowledge base of readers. Apparently, innovation in science and technologies appears to occur far more often in the hard than soft fields, which, thus, makes it clear why natural scientists devote a considerably higher proportion of adverb phrases as code glosses than their social counterparts.

More importantly, the natural sciences are, according to van Noorden's (2015) meta-analysis of 35 million papers published from mid-1980s to 2015, much more interdisciplinary than the social ones. The multidisciplinary nature of such academic texts brings in less specialized audiences. Hence, there is a growing inclination among natural scientists to augment the explicitness of their arguments as a way to render their whole articles more transparent to an increasingly wider readership. In this sense, the denser use of code glosses aids natural scientists in clarifying what they assume may be unfamiliar to the readers outside of their specialist areas.

Additionally, although impersonality still rules in natural scientific writing, the additional function of code glosses to display authority to readers, given the extra information provided in the complement position of a head noun implying writers' larger knowledge base than readers', helps prove scientific competence and establish professional credibility of the writers. This, of course, also appeals to natural scientists for drawing on this metadiscourse strategy more in their work. Similarly, another rhetorical function of reiterating the core messages of the text in the complement position of a head noun through

DOI:10.6814/NCCU202100169

code glosses, as noted above, allows writers to secure readers' better retention of such information, thereby contributing to the wider dissemination of their work.

Taken together, the three functions of code glosses are so significant that even writers in hard domains, though developing a less reader-inclusive relationship between writer and reader, are reluctant to afford the cost of overlooking its use. This also explains why the proportion of code glosses rises to 83.9% in the overall percentage of the adverb phrase use by natural scientists.

In terms of endophoric markers (functioning to direct readers to other parts of the text), the reasons why there are lower frequencies of adverb phrases used as such markers by social scientists can be summarized as the impersonality, the directive/imperativeness and the covert impoliteness embedded in the form of a single adverb (e.g., *above*), as discussed earlier (for details, see Section 5.3.2).

On the other hand, transition markers (e.g., *therefore*) and frame markers (e.g., *firstly*), as pointed out earlier, serve the similar function of rendering the text structure of a research article more explicit, whose uses are denser in the social sciences subcorpus. Granted that the textual organization in soft fields turns out to be more diversified, reiterative and non-linear (for details, see Section 5.3.2), social scientists might find it urgent to guide readers through the text by the use of transition markers and frame markers, while their natural counterparts, who put great stress on the impartiality and linearity of science production and adopt a linear and standardized text structure, tend not to do so. Specifically, scientific findings in hard fields are normally presented in research articles which follow a conventional style and 4-part IMRD format (Hyland, 1998c), which, to a large extent, represents "the standard product of the knowledge manufacturing industry" (Swales, 1987: p. 42). As such, there is no need for writers in the hard knowledge domains to deploy adverb phrases as transition markers or frame markers between a head noun and its complement (a grammatically marked position for these functions), whereas these two metadiscourse devices account for around 15% and 10% of the total adverb phrase use in the subcorpus of social sciences.

DOI:10.6814/NCCU202100169

Finally, as to interpersonal metadiscourse, writers in natural scientific writing express no interest in its employment at all. It seems reasonable since firstly it is not that common for writers in general to place metadiscourse markers in the position between a head noun and its complement, inasmuch as such a post-modifier, though taking the form of just a few words, still might influence the efficient processing of the text by readers. This is because they need to process the intervening grammatical element (an adverb phrase in this case) before reaching the complement clause of the head noun. Moreover, in light of the idiosyncratic mode of knowledge construction and way in which the text is produced in soft fields noted above, social scientists tend to create an interpersonally acceptable persona through attending to the affective expectations of readers, paying them proper respect and even presenting themselves as humble servants of the field. As such, it is not surprising that they seize every single opportunity to construct relations of rapport with their readers, create a sense of disciplinary solidarity and joint endeavor with them, and demonstrate the shared membership of a scientific community with them. All these, of course, involve the placement of interpersonal metadiscourse in the post-modifier position of a head noun to aid in doing so.

On the other hand, the emphasis on building impersonality and objectivity in hard domains and thus the development of a less reader-inclusive relationship between writer and reader seem to exempt natural scientists from grasping each of such chances to make similar rhetorical effort. In other words, once they deem the grammatical position not that appropriate to place these markers, such as the one between a head noun and its complement (having the potential to disrupt the reader's processing of the discourse), they can simply pass up this opportunity and await another. Needless to say, their fields do not especially encourage the use of interpersonal metadiscourse.

5.4.3 Finite dependent clause

Significant interferential statistical differences were also found in the use of finite dependent clauses as post-modification between soft and hard domains (see Table 5.15 below). In general, natural scientists devote all the proportion of finite dependent clauses

DOI:10.6814/NCCU202100169

to attributing stance nouns to abstract entities, whereas, proportionally, soft fields contain a slightly denser use of abstract entities and other humans, averaging around 30%, than

	NSS in CIJA	SSS in CIJA
Abstract entity	15 (100.0)	30 (28.8)
Other human	0 (0.0)	33 (31.7)
Self-sourced	0 (0.0)	19 (18.3)
Self-sourced + Endophoric marker	0 (0.0)	22 (21.2)
Totals	15 (100.0)	104 (100.0)

Table 5.15 Finite dependent clause as post-modification between NSS and SSS in CIJA (% of total)

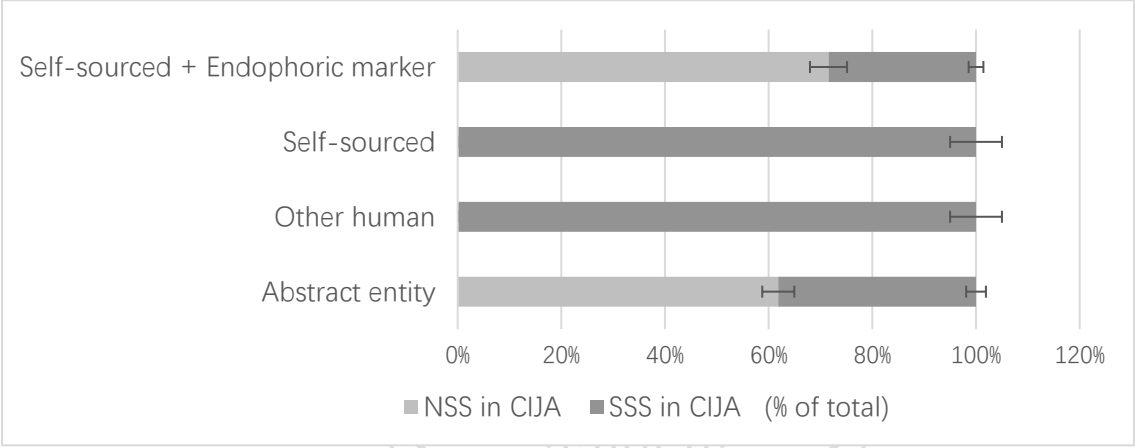


Figure 5.10 Finite dependent clause as post-modification between NSS and SSS in CIJA

that of self-sourced and self-sourced + endophoric markers (approximately 20%) (see Figure 5.10). In terms of abstract entities, knowledge in hard domains, as discussed above (for details, see Section 5.2.1), is constructed principally on empirical evidence and origination of facts through observation and experimentation. This entails the attribution of diverse evidence and facts to their corresponding observation and experimentation as the sources of information, which are, in turn, presented as a wide variety of abstract entities (e.g., *the research*). In this regard, once a head noun is used to refer to such evidence or facts, the placement of finite dependent clauses in its post-modification

DOI:10.6814/NCCU202100169

position can function to aid natural scientists in attributing the information contained in the head noun to different abstract entities as its sources, as illustrated in the following example:

(81) The possibility **which is suggested by the data** [that there is a rise post-operatively in the excretion of these substances] is currently under investigation, both in surgical patients and in stressed non-surgical individuals.

[NSS INT RA38]

Here, the head noun *possibility* is attributed by the following finite dependent clause "which is suggested by the data" to its sources of information *data*, which are, of course, acquired through observation and experimentation and falls into the category of abstract entities. Similarly, in Example (82) below, the finite dependent clause "that underlies the CBB approach" as the post-modifier of the head noun *assumption* attributes it to the abstract entity *CBB approach*, with this *approach* also developed via a series of observations and experiments.

(82) To evaluate the validity of the reported benchmark PRT rate, it is useful to consider the main assumption **that underlies the CBB approach** [that the observed treatment rate will approximate the optimal rate, where there is optimal access and optimal decision-making about the use of RT].

[NSS INT RA40]

As such, there seems to be a far greater need for such attribution to abstract entities of various kinds in hard fields. The deployment of finite dependent clauses as post-modification to fulfil this function offers writers in these fields more grammatical choices to mark the sources of head nouns, thus explaining why they employ finite dependent clauses as post-modification exclusively for the attribution to abstract entities.

Moreover, as to the attribution to other humans, it usually takes the form of an integral citation as noted earlier, i.e., the combination of the last name(s) of the author(s) and the publication year(s), or sometimes the page number(s) in the body of the text, whose focus is, according to Thompson (2001), on the cited authors. However, as Hyland (1998c) points out, it is commonly believed that in natural science articles, writers simply report or

DOI:10.6814/NCCU202100169

describe natural facts mostly without human intervention. From the strict empiricist view of science, scientific discourse is even conceptualized as the neutral descriptive medium wherein natural scientists are generally allowed to assume the role of "a messenger relaying the truth from nature" (Gilbert, 1976: p. 285). In this sense, the employment of finite dependent clauses as integral citations carries the implication of human intervention and is likely to introduce the possible subjectivity and distortions, thereby undermining "the authority of scientific knowledge as built on non-contingent pillars of replication, falsification and induction" (Hyland, 2005b: p. 160).

On the opposite, by reducing their stress on individual researchers and their work through the non-use of integral citations, writers in hard fields not only adhere to the conventions of impersonality in science, but also strengthen the objective ideology that the legitimacy of natural science knowledge is established on "socially invariant criteria" (Hyland, 2004: p. 33). This is further supported by the much lower incidence of citations found in these domains (see, e.g., Hyland, 2000, 2005; Kirkpatrick & Mulligan, 2002; Tang et al., 2020). Additionally, the citations in the hard fields are often in the footnote format, thus substituting numbers for cited authors as in the following example:

> Given evidence **[61]** that legumes differ from non-legumes in their relationships of N_{area} to WUE_i, we also hypothesized that grain legume crops would show stronger relations between these variables than non-legumes.

Taken together, the mode of knowledge construction, the conventions of impersonality, the objective ideology, and the citation format in the natural sciences contribute jointly to the non-use of finite dependent clauses for the attribution to other humans (for details of this pattern of use in soft fields, see Section 5.2.3).

When it comes to the self-sourced use, as the Anglo-Saxon norms of scientific writing entails impersonality, thereby urging writers to conceal themselves in text and replace subjective interpretations by persuasive objectivity, natural science research articles, in particular, have witnessed the scarcity of writers' explicit presence. Moreover, writers from hard domains are, as pointed out earlier, simply "messenger[s] relaying the_truth from

DOI:10.6814/NCCU202100169

nature", cloaked as humble servants to scientific methods (Biber, 2006a; Crosthwaite et al., 2017; Hyland, 2005b; McGrath & Kuteeva, 2012), and exhorted to "strengthen the objectivity of their interpretations and subordinate their own voice to that of unmediated nature" (Hyland, 2001: 216). Also, the overwhelming use of phrasal structures found in the natural sciences, compared to their social counterparts' preference toward clausal forms (Biber & Gray, 2016; Gray, 2015), further contributes to the lower employment of finite dependent clauses in natural science articles. All these thus explain the abandonment of finite dependent clauses for the self-sourced use by writers from hard fields (for details about how finite dependent clauses are used for explicit averrals in the social sciences, see Section 5.2.3).

Last but not least, similarly, writers in international natural scientific journals completely ignore the use of finite dependent clauses to explicitly aver stance nouns, while also referring to other parts of the text (i.e., *self-sourced + endophoric marker*). As noted above, the textual organization in hard knowledge domains, given the impartiality and linearity of science production, turns out to be linear and standardized, thus strictly following a conventional style and 4-part IMRD format most of the time, in a way that represents the standard product of the hard knowledge manufacturing industry. In this sense, the function of endophoric markers to improve the grammatical cohesion and textuality of the text, as previously stated, seems to play a less vital part in the natural sciences than in their social counterparts. Moreover, since natural scientists demonstrate their stronger preference for phrasal over clausal structures, it is reasonable to assume that they are more liable to choose phrasal forms (e.g., adverb phrases) as endophoric markers (e.g., *above*, *below*) instead of finite dependent clauses. As to another function of holding the reader's attention to the unfolding text served by endophoric markers, hard fields, where the conventions of impersonality, objectivity, and the less reader-inclusive relationship (e.g., preference for content-oriented over reader-oriented functions) prevail, tend not to encourage the considerable use of any grammatical forms out of this rhetorical purpose, needless to say in the form of finite dependent clauses. As a result, given natural

DOI:10.6814/NCCU202100169

Feature	Mann-Whitney			Effect size (*r*)
	U	Z	*p*	
Prepositional Phrase	**4982.5**	**-8.696**	**< .001**	***r* = -0.502**
Punctuation	**7152.0**	**-7.009**	**< .001**	***r* = -0.405**

Table 5.16 Lexico-grammatical features of significant inferential statistical differences (in descending order of effect size *r* value) between NSS and SSS in CCLJA

scientists paying no attention to the use of finite dependent clauses for self-sourced stance-making as just discussed above, it is not surprising that they make the similar choice of entirely abandoning such clauses for overt averrals and concurrently as endophoric markers (for details of finite dependent clauses used for the *self-sourced + endophoric marker* function in soft fields, see Section 5.2.3).

5.5 Explanations for different use between local writers in the natural and social sciences

Writers from local journals in the hard and soft fields express quite different preferences for using grammatical elements as post-modification to perform linguistic and extra-linguistic functions. The statistical differences are particularly noticeable for two lexico-grammatical features in the two subcorpora – prepositional phrases and punctuation (see Table 5.16). In this section, the lexico-grammatical feature *prepositional phrase*, given its highest statistical significance (Effect size: *r* = -0.502), will be analyzed first.

5.5.1 Prepositional phrase

The deployment of prepositional phrases for the attribution to abstract entities, proportionately, tends to dominate in the local hard domains, with the greatest density of it as well in the local soft fields (see Table 5.17 and Figure 5.11 below). This is not surprising in the nature sciences where the fundamental need for the attribution to abstract entities of various kinds (e.g., data from experiments), as previously discussed, entails as many grammatical choices (e.g., prepositional phrases) as possible for hard scientists to mark abstract entities as the sources of head nouns. Moreover, the reasons why local social

DOI:10.6814/NCCU202100169

	NSS in CCLJA	SSS in CCLJA
Self-sourced	0 (0.0)	5 (1.1)
Other human	1 (1.5)	166 (35.5)
Abstract entity	54 (84.4)	209 (44.8)
Qualifier + subject	9 (14.1)	87 (18.6)
Totals	64 (100.0)	467 (100.0)

Table 5.17 Prepositional phrase as post-modification between NSS and SSS in CCLJA (% of total)

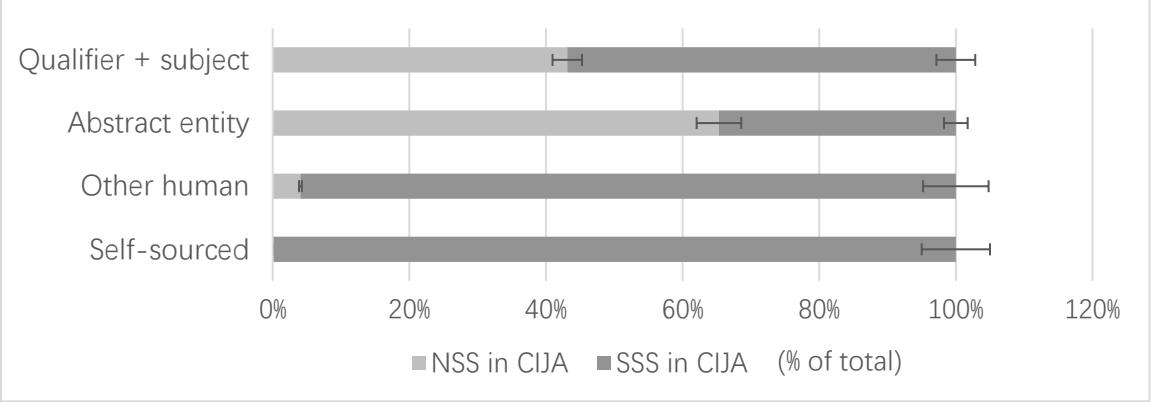


Figure 5.11 Prepositional phrase as post-modification between NSS and SSS in CCLJA

scientists tend to draw on most of the prepositional phrases as post-modification for attributing to abstract entities are more complicated.

In the first place, it may, to a large extent, pertain to local soft science writers' neglect of the roles that the function of attributing to other humans assumes in their post-modification use. Specifically, the use of prepositional phrases for the attribution to other humans normally takes the form of integral citations in the soft knowledge domains, a practice of incorporating others' names into the body of the sentence instead of placing them in parentheses or a footnote (Gray, 2015; Maher, 2015; Thompson, 2001, 2012). Besides, writers in soft fields tend to build arguments and knowledge based on their personal interpretations and negotiations with readers, whereby new arguments and knowledge are not advanced merely from writers' own thoughts and ideas but rather

DOI:10.6814/NCCU202100169

"follow[] altogether more reiterative and recursive routes as writers retrace others' steps and revisit previously explored features of a broad landscape" (Hyland, 2004: p. 31). In addition to this, since readers are not presumed to have the same interpretative knowledge, writers need to establish a context through citations. As such, the use of integral citations aids writers in developing a discursive and contextual framework within which writers can negotiate agreement of their interpretations of texts with readers, on the premise that both sides possess similar disciplinary understandings. More importantly, this pattern of use allows writers to attribute ideas and insights to prominent disciplinary figures (e.g., *Rod Ellis* in the field of second language acquisition), thus bringing themselves into an alignment with a particular scholarly community. In so doing, their new interpretations and evaluations, given having been embedded in the literature of the community, can then exhibit their "relevance, importance and the credentials of the writer" (Hyland, 2005b: p. 158).

Taken together, it seems reasonable to assume that local social scientists are liable to be unaware of all the discursive and interpretative nature of the soft knowledge domains in English research article writing, or in other words, their rhetorical awareness of this pattern of use (i.e., *other human*) needs to be improved, which, according to Mu (2020), usually takes a long time for non-anglophone writers. Therefore, such rhetorical unawareness might give rise to lowering the proportion of prepositional phrases used for the function of other humans, which is, then, surpassed by that of abstract entities.

Secondly, there seems to be general consensus among scholars within the field of academic writing that detecting a major problem in others' work is one of the main ways that justify their research to readers in the soft domains (e.g., Becher & Trowler, 2001; Gilbert & Mulkay, 1984; Hyland, 2004; Jiang, 2017; Swales, 1990). Also, as Bloch and Chi's (1995) study suggests, there was almost no difference in the frequencies of criticizing other studies between Chinese writers and their Anglo-American counterparts in the social sciences. Or, in their own terms, a similar "critical edge" (p. 249) also exists in the texts written by Chinese scholars, given the emphasis on *thinking* rather than merely *memorizing*

DOI:10.6814/NCCU202100169

in the teachings of Confucius encouraging critical reasoning as well. Nevertheless, Chinese writers' preference for rhetorical strategies of indirectness helps shape their more cautious and implicit styles when critically assessing the value of others' work (Kaplan, 1972; Liu & Du, 2018; Scollon, 2001; Scollon & Scollon, 1995).

In this sense, attributing the sources of stance nouns to abstract entities (e.g., the results or research methods) instead of referring them directly to other scholars aids in concealing the sources of the evaluation, thereby mitigating threats to their face and demonstrating writers' goodwill gesture to maintain the rapport with them. This also explains the denser use of prepositional phrases for the function of abstract entities than that of other humans in the soft knowledge domains.

Additionally, Chinese writers in the soft knowledge domains are believed to be more "tied to the past" and thus more inclined to "rely on older, classic texts" in L1 Chinese writing (Bloch & Chi, 1995: p. 248). They are also encouraged to cite extensively from the works of the sages (*yǐn jīng jù diǎn*), which is, to a large extent, deemed "a willingness to respect authorities and to accept traditional values, social norms, and group ideologies, and [] a desire to be polite" (Cai, 1999: p. 283). These underlying cultural beliefs and discursive practices therefore seem to restrict the scope of others' work considered appropriate to be cited, since, most importantly, the cited authors need to be regarded as the respected authorities in their specialist fields. Also, they are expected to share culture-specific values, norms and ideologies with text writers, and their work had better be time-honored and classic. As such, the limited amount of qualified cited authors is likely to reduce the occurrences of integral citations, which may, then, cause the effect of negative transfer on their English research article writing. This might help explain the less use of the attribution to other humans in the local soft fields.

Lastly, Chinese writers tend to directly quote others' work, but without giving credit to the original authors (see, e.g., Cai, 1999; Zhang, 2011), especially in L1 Chinese writing in mainland China, which is, however, liable to be considered plagiarism in English academic writing. This is because people in mainland China often view knowledge as

DOI:10.6814/NCCU202100169

commonly shared heritage rather than private intellectual property for certain historical and socio-political reasons. More importantly, although cited and/or citing references in English writing styles and formats (e.g., APA style) have been widely used in local English journals, the submission guidelines on reference styles are usually not specific and systematic in local journals, particularly natural scientific ones in mainland China, with the manuscripts not conforming to these guidelines still accepted and published, mainly due to the shortage of highly qualified English editors (see, e.g., Ding & Miao, 2020; Qiang & Hua, 2010; Yin, 2016). This, of course, decreases the number of integral citations, of which the attribution to other humans usually takes the form. In addition, the citations in the local hard fields are often in the footnote format, in a way that substitutes numbers for cited authors. All these factors therefore cause the percentage of prepositional phrases devoted to the function of other humans to decline, in particular in the hard fields, where the proportion falls even to 1.5%.

In terms of the self-sourced use, writers from both local natural and social science journals express little or no enthusiasm for it. This is mainly due to the sociocultural ideologies and Chinese people's general view on the development of arguments. Specifically, argumentative writing in the Chinese language aims to *tell* readers "what the world should be" rather than *show* to them "how the world is" as in its English counterpart (Coffin, 2004: p. 231). As such, the native Chinese writers seem to be exempt from the focus on convincing readers of their viewpoints in their texts, as they are just telling the truth about what the world really is through analogies, maxims and citations from early scholars and researchers' work. In fact, the deeply held sociocognitive belief that "verbal debate and argumentation are not meaningful tools for understanding truth and reality" (Peng & Nisbett, 1999: p. 15) is, as noted earlier, highly characteristic of Chinese cultural practices. To put it differently, all truth is considered self-evident thus in no need of arguments (Bodde, 1991; Hu & Cao, 2011; Yang, 2013). In this sense, knowledge is supposed to be accumulated on the basis of previous convictions and experience (Yang, 2013), and authoritative knowledge (i.e., the resource for creating a communal way of

DOI:10.6814/NCCU202100169

viewing the world) is profoundly respected (Davis-Floyd & Sargent, 1997; Nakayama & Dusenbury, 1984; Tweed & Lehman, 2002). Holding these conventional values of respecting authorities and accepting traditions, Chinese writers frequently draw on impersonal forms and distance themselves from the arguments as a way to assume the role of a "truth-teller". This thus minimizes the visibility of the writer and maximally decreases the frequency of prepositional phrases as post-modification for overt averrals (even 0% of the self-sourced use by local natural scientists, see Table 5.17 above).

Although Zhang & Zhan (2020) point out that certain Chinese graduate students born in the 1980s increase their authorial presence in academic writing because of the changing sociocultural context in mainland China, particularly China's economic prosperity and one-child policy, the authorship of the local journals in the social sciences included in the subcorpus are mainly writers who grew up before the reform and open-up policy of China (before the 1980s) and are thus more subject to Chinese traditional values. More broadly, since maintaining social stability and harmony is deeply embedded in Confucian values and ideology, individuals in mainland China are usually made responsible for such maintenance. For a long period of time in Chinese history, self-expression of one's personal views was considered to result in conflicts, rifts, and even cataclysms, thereby imposing harmful effects on both self-expressing individuals and social harmony. However, since Reform and Opening up started more than 40 years ago, Chinese society has been much more open to the presentation of personal voice. Notwithstanding this, in terms of political issues and policy making, people still refrain themselves from voicing clear personal views on these subjects whose central role in the soft knowledge domains, though triggering numerous opportunities for self-expression, is still not able to aid Chinese social scientists in overtly referring more sources of information to themselves. This is supported by only 1.1% of prepositional phrases as post-modification used by them for the self-sourced function (see Table 5.17 above).

Last but not least, proportionately, local natural and social scientists make similar less use of prepositional phrases for the *qualifier + subject* function. Local writers' reluctance

DOI:10.6814/NCCU202100169

to employ it, as discussed above, is largely attributed to the grammatical complexity of this particular construction, namely, that the prepositional phrase functions simultaneously as the qualifier of its preceding stance noun and as the subject of its following proposition (e.g., the admission *for ACEs* to compare the periods before and after the ban). Besides this, there now exists a broad consensus among scholars in this line of research (see, e.g., Biber, 2006a; Biber & Gray, 2013, 2016; Egbert, 2015; Gray, 2015; Staples et al., 2016) that natural scientists tend to favor the phrasal complexity, a grammatical phenomenon where clauses are maximally compressed into complex phrases, whereas the clausal complexity, or the *T-unit* (i.e., a main clause and all associated dependent clauses), prevails in the soft fields. This is because the incorporation of multiple dependent clauses, normally with multiple levels of embedding, is usually drawn on by social scientists to render their argumentation elaborated and explicit in a way that fits the mode of knowledge construction based on plausible reasoning and explicitly interpretative persuasion in such domains.

On the other hand, natural science writers, given the linear and standardized, almost shorthand, argument structure shaped by the knowledge construction on hard facts and empirical evidence through observation and experimentation in these domains, opt not to bother themselves with elaborated sentence structures and clausal complexity but to embrace the phrasal complexity, despite its associated loss in elaboration and explicitness of meaning. In this sense, local natural scientists should have devoted more prepositional phrase proportionally to the qualifier + subject use, given its compact, integrated and phrasal structure. Nevertheless, their relatively lower levels of proficiency in English seem to limit the good access to this pattern of use, since a clear understanding of the relationship

	NSS RAs in CCLJA	SSS RAs in CCLJA
Single prepositional phrase	9 (100.0)	75 (86.2)
Complex prepositional phrase	0 (0.0)	12 (13.8)
Totals	9 (100.0)	87 (100.0)

Table 5.18 Types of prepositional phrase used between NSS and SSS in CCLJA

DOI:10.6814/NCCU202100169

among stance noun, prepositional phrase as post-modification, and complement clause, in particular the dual role of the post-modifier (i.e., the prepositional phrase in this case), is needed. This is further supported by their placement of just the single rather than complex prepositional phrases in post-modification position (see Table 5.18), indicating their attempts to decrease the grammatical difficulty in applying this function.

5.5.2 Punctuation

There is the significant statistical difference indicated by the Pearson *r* value in the use of punctuation as post-modification between the writers from local journals of the natural and social sciences. Overall, local social scientists exhibit strongest preference for commas, followed by dashes, with colons used almost three times less frequently than commas, whereas dashes and commas are equally favored by their natural counterparts, who, on the other hand, demonstrate complete disregard for colons (see Table 5.19 and Figure 5.12).

	NSS in CCLJA	SSS in CCLJA
Comma (,)	6 (50.0)	58 (46.4)
Colon (:)	0 (0.0)	21 (16.8)
Dash (—)	6 (50.0)	46 (36.8)
Totals	12 (100.0)	125 (100.0)

Table 5.19 Types of punctuation used between NSS and SSS in CCLJA (% of total)

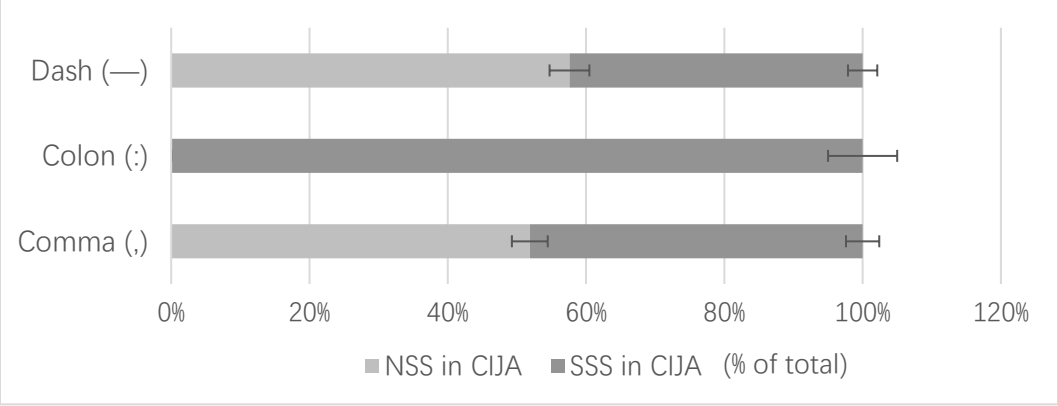


Figure 5.12 Punctuation as post-modification between NSS and SSS in CCLJA

DOI:10.6814/NCCU202100169

Firstly, when it comes to colons as post-modification, they usually function to connect a head noun and its complement clauses, with the aim of signaling the explanatory relationship between the information prior to and after this punctuation (Biber & Gray, 2016). In so doing, writers in fact maintain a reader-oriented attitude and take "the responsibility to make clear and well-organized statements" (Hinds, 1987: p. 143). As the complement clauses following the colon can be of large grammatical complexity (e.g., in the form of coordinated clauses), the readers, especially non-expert readers, may be unfamiliar with how academic texts present the content of this kind. Thus, they are likely to find it difficult or problematic to figure out this meaning relation without the cue given by a colon. However, in the Chinese language cultures, readers rather than writers are made primarily responsible for effective communication in the way that "writers compliment their readers by not spelling everything out, while readers are said to savor hints and nuances" in order to dig out meanings (Hyland, 2005b: p. 116). Granted that "basically the L2 writer is writing from his or her own familiar culture" (Hyland, 2003: p. 47), it is not surprising that local natural scientists tend to overlook this function of facilitating readers through the use of colons as post-modification.

One may argue that since local writers from social science journals share the same reader-responsible cultures with their natural counterparts, why they still attempt to draw on colons for the purpose of explicitly guiding readers through the text. Actually, this cultural tendency of expecting the reader to invest more in the text is likely to be in tension with the mode of knowledge construction underlying the soft domains. Specifically, in soft fields, knowledge is constructed on plausible reasoning and explicitly interpretative persuasion, which, in turn, entails a co-production of the work by the writer and by members of the audience to which it is directed. In this sense, the written work is only finished when writers have managed to bring readers into alignment with their views and establish a common ground between the two parties, or in other words, when the reasoning has been accepted and the persuasion has succeeded. As such, as co-producers of the text, social science writers, undoubtedly, would seize every opportunity to facilitate the co-

DOI:10.6814/NCCU202100169

producing process undergone by readers. This, of course, includes using the colon to give them the cue to the explanatory relationship between the information prior to and after it. The existence of such a tension between writers' native cultures and disciplinary characteristics is further supported by the fact that the proportion of colons as post-modification turns out to be the lowest one, compared to that of commas and dashes.

Clearly, the reluctance of local natural scientists to draw on long and complex coordinated clauses also contributes to their non-use of colons. In addition, their rhetorical unawareness of using colons to connect either a head noun and a dependent *wh*-interrogative clause, or a substantive head noun and a simple complement clause, as noted above, results further in their total neglect of colons as post-modification.

As to the use of commas, in the local subcorpora they only perform the function of showing an identity meaning relationship between a head noun and its nominal complement clause. Since the 20th century, as previously stated, appositive noun phrases of considerable length and complexity have prevailed in scholarly writing for conveying the message with an identity relationship to the stance noun. This occurs in both the natural and social sciences alike, which, in turn, entails an increasing use of nominal complement clauses to replace for those long and complex noun phrases, granted that a clausal structure provides far more grammatical positions than a phrasal one to include as many meaning elements as possible. One may argue that writers from the hard knowledge domains are, given their far stronger preferences for phrasal complexity, much less likely to employ a clausal element (e.g., a *that*-complement clause) to specify or explain the head noun than their social counterparts. Therefore, why do they devote a half proportion of their punctuation use to connecting a head noun and its complement clause with a comma, which is, surprisingly, even slightly higher than that in the local social sciences subcorpus (50.0 vs. 46.4)?

The corpus-informed investigation of lexical bundles across genres, registers, and speakers/writers can help explain this. In early literature, Biber et al. (1999) note that the major structural difference between phrasal and clausal complexity can be revealed through

DOI:10.6814/NCCU202100169

exploring how most lexical bundles in a text bridge two structural units. In their definition, a lexical bundle starts at the boundary of a phrase or clause, and the last word of the bundle is the first element of the following structural unit (e.g., *the situation of the*) which, therefore, determines whether it bridges either two phrases or clauses. Following this, corpus-based studies on specialized registers and writers with different backgrounds demonstrate significant differences in the usage patterns of such phrasal and clausal chunks between conversation and academic prose (e.g., Biber et al.,1999; Biber & Conrad, 1999), university classroom teaching and textbooks (Biber et al., 2004), expert and novice writers (e.g., Chen & Baker, 2010; Cortes, 2002, 2004), and native and non-native English writers (e.g., Biber & Conrad, 2014; Hyland, 2008). The results in this line of research show that the more informal the register is, the more clausal the lexical bundles are, while expert or native English writers, compared to their novice or non-native counterparts, tend to draw on far more phrasal chunks in academic writing.

More recently, Pan et al. (2016) made structural and functional comparisons of lexical bundles in the English research articles written by L1 English and Chinese expert writers. Their findings reveal that Chinese writers favor clausal over phrasal chunks, or in other words, they are in the process of transforming from the more use of clausal chunks into that of phrasal chunks. This can be attributed to their unawareness of the nature of phrasal chunks being the unique grammatical characteristic of advanced academic writing in English, and to their lack of access to the textbooks and systematic trainings in English academic writing, particularly in the structural and functional properties of phrasal chunks. In this sense, given their denser use of clausal structures, it is not surprising that local natural and social scientists as L1 Chinese writers are more inclined to draw on simple nominal clauses rather than complex noun phrases as a way to provide the specific identity or additional descriptive information about the head noun, which, indubitably, leads to the increased incidence of comma in their work to signal such an identity relationship.

More importantly, there are usually less textbooks or trainings in academic writing in hard domains in China, since the effective curriculum design or instructional methods for

DOI:10.6814/NCCU202100169

EFL academic writing are conspicuously missing from most of the course descriptions or handbooks in the department or college of the natural sciences at Chinese universities. Granted that the shortage of textbooks and trainings in English academic writing contributes greatly to the more use of clausal than phrasal phrases in Chinese writers' English research articles, it is reasonable to assume that local natural scientists are more likely to substitute simple nominal clauses for complex noun phrases to help establish an identity relationship signaled by a comma than their social counterparts. This thus slightly increases the proportion of commas in their total use of punctuation as post-modification.

Last but not least, dash, similar to comma, is usually used to connect a head noun and its expository apposition of various forms which, of course, include nominal complement clauses, while it is also "the least formal" way of showing apposition (Seely, 2007: p. 84). As such, the first reason why proportionately, local natural scientific writers make more use of this punctuation mark as post-modification than their social counterparts can be attributed to their rhetorical unawareness of the informality involved in this pattern of use, which, then, probably derives from their lack of textbooks and trainings in English academic writing. More importantly, this also pertains to the Chinese punctuation system. Specifically, in the Chinese language, dashes tend to be interchangeably used with colons (Li, 2018), which might lead to the effect of negative transfer (see, e.g., Benson, 2002; Ellis, 2008) on local writers' discrimination in the levels of formality between these two marks. As a result, they are likely to treat dashes and colons as being at the same formality level in English.

Moreover, granted that dash and colon serve similar functions and take the same forms in English and the Chinese language, English language teachers in China are liable to overlook the teaching of the differences in formality between these two punctuation marks in English and Chinese. Hence, they fail to counteract the influence of negative transfer exerted by the Chinese punctuation system upon the learners' discrimination between the two marks in English.

DOI:10.6814/NCCU202100169

Besides, Chinese writers are more liable to replace colons with dashes, but not vice versa, not only because dash marks a clearly stronger break from the immediate co-text and allows a richer variety of grammatical elements to follow it than colon in the English punctuation system (Huddleston & Pullum, 2002), but also because in the Chinese language, dash often introduces something which further develops or exemplifies what has been written before. This enables it to be omitted from a text without changing its meaning, thus rendering its use more grammatically flexible (e.g., easier to insert it into a clause), not to mention its function of emphasizing the text introduced. In contrast, colon only helps list items which are rather lengthy and therefore cannot be deleted from the text (see, e.g., Lin, 2000; Zhong, 2006).

In fact, although dashes and colons are interchangeable in Chinese, there are still noticeable differences in their uses (see, e.g., Lin, 2000). As such, local natural scientists' less access to the textbooks and trainings in the rules of Chinese grammar especially for punctuation marks than their social counterparts is more likely to lead them to form a vague impression. Namely, dash and colon are similar punctuation markers, but the former suits most of the grammatical and rhetorical purposes, which, in turn, increases their frequencies of replacing colons with dashes in their Chinese scholarly writing. In this regard, it is reasonable to assume that every time local writers in hard domains are faced with the choice between dash and colon, they would not hesitate to opt for the former, and that given the influence of negative transfer from the Chinese language, they are most likely to follow the same pattern in their English academic written discourse. Taken together, local natural scientists' lack of knowledge about both English and Chinese punctuation systems results in their denser use of dashes than their social counterparts.

5.6 Summary and comparison of explanations for cross-disciplinary and cross-cultural differences with prior research

To summarize, through the explanatory model of disciplinary and cultural use of post-modification in the *Noun Complement* construction, this research provides a comprehensive picture of the linguistic, rhetorical, relational, and socio-cultural

DOI:10.6814/NCCU202100169

dimensions of why the target structure functions differently in the work of writers from contrasting disciplinary fields and linguistic-cultural backgrounds. Specifically, the reasons behind the cross-cultural differences were analyzed based on the following seven major factors in the above four dimensions: namely, (a) idiosyncratic features of the semantic, syntactic and punctuation systems in the Chinese language (e.g., non-existence of the *Noun Complement* construction), (b) Chinese EFL writers' language competence in English academic writing (lower levels of proficiency in English) in the linguistic dimension, (c) traditions of Chinese rhetoric (e.g., obsession with flowery language and big words to demonstrate one's literary skills), (d) Chinese EFL writers' academic competence in English academic writing (e.g., rhetorical unawareness of the underlying functions of attitude markers) in the rhetorical dimension, (e) notions of writer-responsibility in the West versus reader-responsibility in the East (e.g., lifting writers' responsibility to guide readers through the text in the latter) in the relational dimension, (f) Chinese cultural practices rooted in Confucian and Taoist traditions (e.g., discouragement of self-expression of one's personal views, particularly in terms of political issues and policy making), and (g) influence of social and political movements in Western cultures (e.g., post-modern and feminist movements less affecting Chinese EFL writers) in the socio-cultural dimension.

In so doing, this study has developed a more complete understanding of various influences on the use of the target structure between writers of different languages and cultures than previous studies, whose prime focus is narrowly on accusing Chinese EFL writers of their low levels of proficiency in academic English (e.g., limited writing lexicon, mixed use of spoken and written forms, unawareness of conventions of English academic discourse, etc.), or English language teachers in China of their improper instruction (e.g., encouraging students to use "generalized and holistic" words).

Similarly, the underlying motivations for the cross-disciplinary differences in the use of the target structure were identified according to the contributing factors in the same four dimensions: (a) linguistic conventions in the genre of research articles (e.g., the extent to

DOI:10.6814/NCCU202100169

which impersonal forms are used), (b) nature of the grammatical complexity differences (preferences for phrasal complexity in hard domains vs. clausal complexity in soft domains) in the linguistic dimension, (c) modes of knowledge construction (reliance on hard facts and empirical evidence collected through experimentation and observations vs. plausible reasoning and explicitly interpretative persuasion), (d) disciplinary rhetorical traditions (e.g., the extent to which directness is adhered to) in the rhetorical dimension, (e) interpersonal writer-reader relationship (e.g., preference for content-oriented over reader-oriented functions in hard sciences) in the relational dimension, (f) developmental differences in scientific research (e.g., the globally competitive environment for scientific and technological innovation leading to far more complex research in hard than soft fields), and (g) impacts of social and political movements in Western cultures (slight vs. deep influence in the natural and social sciences) in the social-cultural dimension. As a result, the present study more systematically displays the plausible reasons behind the cross-disciplinary use of the target structure than other studies conducted in this vein, most of which, however, attribute disciplinary differences exclusively to discipline-specific factors (e.g., the contrasting ways that knowledge is built between the natural and social sciences).

Taken together, the comparisons of cross-disciplinary and cross-cultural differences grounded on the four dimensions of the explanatory model in the present study provide an integrated reason analysis of post-modification in the *Noun Complement* construction. By identifying the reasons behind such differences through examining the linguistic, rhetorical, relational, and socio-cultural aspects of how the target structure works, this research in fact treats any disciplinary field or culture as situated in a wider social world, which, in turn, produces a rich variety of motivations and enabling factors that underlie the surface variations. This, however, is unexpected by previous researchers, given their analyses of disciplinary and cultural use within the limited scopes of the disciplines and the ESL/EFL writers per se. More importantly, the way in which the cross-disciplinary and cross-cultural discussion of the reasons behind these variations is conducted in this study actually help explain why post-modification in the *Noun Complement* construction is used widely and

DOI:10.6814/NCCU202100169

in quite various ways. Namely, writers from different disciplinary fields and cultures possess their own idiosyncratic linguistic, rhetorical, relational, and socio-cultural ways of projecting their stances through any linguistic form, even the one placed in the grammatically marked position as in the case of the present study.

5.7 Comparison of lexico-grammatical features with prior research

Table 5.20 compares the lexico-grammatical features of post-modification in the *Noun Complement* construction with that of general noun phrase (NP) post-modification, and pre-modification in the *Noun Complement* construction identified in prior research in this vein. Overall, the use of post-modification contains the largest number of lexico-grammatical features identified (8 total), which, more importantly, covers all such features of the general NP post-modification and pre-modification use. Specifically, compared to general NP post-modifiers, post-modification, in terms of phrasal structures, has been found to take the additional forms of adverb, adjective or noun phrases. As to the clausal forms, post-modification also outperforms general NP post-modifiers with the addition of *to-infinitive* clause in the category of non-finite clauses and adverbial clause in the category of finite dependent clauses. Additionally, punctuation and extraposition are only identified as lexico-grammatical features of post-modification.

Lexico-grammatical features		
Post-modification	**General NP post-modifiers**	**Pre-modification**
Prepositional phrase	Relative clause	Adjective
Adverb phrase	*Ing*-clause	Possessive
Adjective phrase	*Ed*-clause	
Noun phrase	Prepositional phrase	
Non-finite clause	*Of*-phrase	
Finite dependent clause		
Punctuation		
Extraposition		

Table 5.20 Lexico-grammatical features of post-modification, general noun phrase modifiers, and pre-modification

DOI:10.6814/NCCU202100169

In the cases of pre-modification use, the adjective forms are normally used as single adjectives in attributive position, whereas they can take the form of either single adjectives or adjective phrases in the use of post-modification. Besides, the possessive form in pre-modification is actually equal to several feature categories of post-modification. To illustrate, possessive nouns as pre-modifiers (e.g., *Biber's* theory) are equivalent to *of*-phrases (the theory *of Biber*), *ed*-clauses (the theory *developed by Biber*), or relative clauses (the theory *that Biber developed*) as post-modifiers. The examples of these lexico-grammatical features haven been presented in Chapters 2 and 4, thus omitted in this section.

To conclude, the present study expands the range of investigations into the feature categories of post-nominal modification, particularly that in the *Noun Complement* construction. The stereotypical view previous researchers have held based on native-speaker or ESL/EFL-learner intuitions (given part of researchers as ESL/EFL learners themselves) for post-modification in the *Noun Complement* construction have generally assumed that its lexico-grammatical features are severely limited, which, thus, discourages them from taking a close look at this post-nominal structure. In contrast, based on these rich and varied grammatical forms from phrases to more complicated clausal structures, this dissertation has further witnessed a much broader variety of linguistic and extra-linguistic roles assumed by such forms than any other study reported to date in *Noun Complement* construction research.

5.8 Comparison of functions with prior research

This dissertation provides a comprehensive picture of the linguistic and extra-linguistic aspects of how post-modification in the *Noun Complement* construction functions in the work of writers from contrasting disciplinary domains and cultures (see Table 5.21). In terms of linguistic functions, they were identified based on whether the lexico-grammatical features with statistical significance (effect size (r) $\geqq \pm.300$) are used for stance noun and its complement clause at *Noun Complement* construction level, or for information structure and cohesion at discoursal level. Specifically, these features of post-modification can be employed to introduce or signal identity relationships, elaborative interpretations, complex

DOI:10.6814/NCCU202100169

Post-modification	
Language intrinsic functions	**Extra-linguistic functions**
Identity relationship	Interpersonal metadiscourse markers
Elaborative interpretation	Hedge
Complex grammatical structure	Booster
Dependent wh-clauses	Attitude marker
Substantive expression	Abstract entity
Qualifier + subject	Other human
Interactive metadiscourse marker	Self-sourced
Transition marker	Informality
Frame marker	Strategic effort
Endophoric marker	Self-sourced + endophoric marker
Code gloss	
Information principle	
Principle of end-weight	

Table 5.21 Functions of lexico-grammatical features with statistical significance in the use of post-modification

grammatical structures, dependent *wh*-clauses, and substantive expressions, as the qualifiers of the stance noun and concurrently the subjects of its complement, transition markers, frame markers, endophoric markers, code glosses, or for applying the information principle and the principle of end-weight. As to the extra-linguistic functions, they were ascertained according to the roles the target features play that concern the writer, the relationship between writer and reader, the production and comprehension demands of research article writing, and the cultural preferences. In this regard, post-modification is able to serve as hedges, boosters, attitude markers, to attribute stance nouns to abstract entities, other humans, or writers themselves, to display informality and launch strategic effort, or even to both overtly aver the stance noun and refer it to other parts of the text.

By comparison, in the previous studies into the general NP post-modification and the

DOI:10.6814/NCCU202100169

Functions	
General NP post-modifiers	**Pre-modification**
Expressing concrete/locative meanings:	First-person possessives/Overly averred
Genitive meaning relationships	Attributive possessives/Abstract entity
Locative meanings	Scholarly citation/Other human
Additional descriptive information	Offering additional chances for stance-making
Expressing abstract meanings:	Showing writers' attitudinal feelings
Identifying topical domains	Establishing an authorial presence
Marking semantic patients	
Describing abstract states	
Identifying purposes of entities	
Establishing evidentiary bases	

Table 5.22 Functions identified in prior research into general NP post-modifiers and pre-modification

pre-modification in the *Noun Complement* construction (see Chapter 2, Section 2.5.1 and 2.5.2), general NP post-modifying devices are usually found either to communicate a number of different meaning relationships with their head nouns (e.g., expressing genitive, locative, or identity meaning relationships), or to convey varying semantic associations with them (e.g., identifying topical domains, describing abstract states) (see Table 5.22). The functions of pre-modification in the *Noun Complement* construction, otherwise, are largely centered around the ownership of the stance noun (e.g., attributing to other researchers), with certain research touching upon the extra-linguistic roles played by pre-modification structures (e.g., showing writers' attitudinal feelings). Nevertheless, such roles are either too general (e.g., offering additional chances for stance-making) in that they are not categorized and specified (e.g., what kinds of chances), or concerned with negative examples of how pre-modifiers are functioning (showing writers' attitudinal feelings or establishing an authorial presence) that need to be avoided in academic writing.

DOI:10.6814/NCCU202100169

To conclude, the rich and varied grammatical forms from phrases to more complicated clausal structures of post-modification in the *Noun Complement* construction render this linguistic structure capable of playing a wide variety of linguistic and extra-linguistic roles in text through such forms. This is because phrasal and clausal structures tend to occupy an extensive range of positions in a clause, thereby incorporating more meaning elements into it as well. In other words, such forms of post-modification provide writers more grammatical slots to fill meaning elements, which can, then, be used to construct and perform linguistic and extra-linguistic functions in myriad and various ways. Thus, it is reasonable to propose here that it is worth looking closely at any grammatical component in a stance-making structure, even a grammatically-marked and stereotypically-associated one (like the target structure of this research), to explore its research value.

DOI:10.6814/NCCU202100169

CHAPTER 6

Conclusion

6.1 Overview

This dissertation is the first to employ a corpus-based and mixed-methods approach to examine the use of post-modification in the *Noun Complement* construction by published academic writers of English-medium journal articles. As such, it makes a number of contributions to studies of *Noun Complement* construction. First, the present study is unique in that it applies the inferential statistical methods for the first time in *Noun Complement* construction research to identify the variances in the lexico-grammatical features of the target structure between disciplinary fields and writers of different native languages and cultures. Second, it also reveals the cross-disciplinary and cross-cultural differences in the functions served by the features with significant inferential statistical differences, and interprets the findings qualitatively through the self-constructed analytical framework for functional analysis. Another innovation is that it explored the underlying motivations and enabling factors behind disciplinary and cultural use of the target structure from the linguistic, rhetorical, relational, and socio-cultural dimensions of the explanatory model developed in this study, thus the discussion of such reasons beyond the limited scope of the disciplines and writers per se in prior research. Finally, it contributes to existing knowledge of research in this vein by using the optimized combination of regular expression queries, based on "a well-established, fairly standard and extremely powerful search syntax" (McEnery & Hardie, 2011: p. 255), for extracting all the cases of nominal complement clauses. This methodology enables researchers to extract cases of the target structure more efficiently and accurately than the method adopted in most of the prior studies into *Noun Complement* construction. Namely, it searches texts for the words tagged as *noun* and *plural-noun* before eliminating irrelevant cases.

In Chapter 1, the importance of investigating the use of post-modification in the *Noun Complement* construction has been attributed to (a) its popularity in the stance repertoire of academic writers in various disciplinary domains and with different first languages and

DOI:10.6814/NCCU202100169

cultural backgrounds, contradicting the stereotypical view of it based on unreliable intuitions, (b) its rich and varied grammatical forms from phrases to more complicated clausal structures, (c) its capability for playing a wide variety of linguistic and extra-linguistic roles in text through such forms, (d) the particular interest of uncovering the reasons behind academic writers' insistence on deploying this grammatically marked structure that syntactically and semantically breaks apart head noun and its complement clause, and (e) a very real dearth of empirical studies into this structure. The theoretical framework and analytical methods for studying post-modifications in *Noun Complement* construction was then developed and refined in Chapter 2 based on published studies on the *Noun Complement* structure as a nominal stance construction, Biber and his associates' typological theories, the analytical approach taken in the line of research into pre-modification, and theories about functions of language. Next, a new corpus was designed and built to include research articles balanced between two branches of science (i.e., natural vs. social sciences) and two groups of writers (i.e., contributors of English-medium journals based in the UK or US vs. in Taiwan or mainland China) (see Chapter 3). Additionally, Chapter 4 presents the results of the cross-disciplinary and cross-cultural differences in the frequencies, forms and functions of post-modification in *Noun Complement* construction. Finally, the underlying motivations behind these variations were discussed in Chapter 5, which is then followed by the comparison of findings between this dissertation and previous studies.

The goals of this chapter are to (a) recap and further discuss the findings given in the prior chapters, while also comparing them with what have been found in other studies where necessary (Section 6.2), (b) discuss certain implications of these results (Section 6.3), and (c) conclude with a discussion of the limitations of the study and directions for future studies (Section 6.4).

6.2 Summary of findings

This section will sum up the key findings of the present study. First of all, the use of post-modification in the *Noun Complement* construction by writers from contrasting

DOI:10.6814/NCCU202100169

disciplinary fields and cultures contains the greatest variety of lexico-grammatical features ever identified in previous studies in this vein. Specifically, in the category of phrases, it includes prepositional, adverb, adjective, and noun phrases. The category of clauses comprises non-finite clauses, including *ing*-clauses, *ed*-clauses, and *to-infinitive* clauses, and finite dependent clauses, composed of relative and adverbial clauses. The other two categories of punctuation and extraposition are also incorporated as the lexico-grammatical features into the use of the target structure. As such, the identification of these rich and varied grammatical structures from phrases to more complicated clausal forms contributes to broadening the range of analyses into the feature categories of post-nominal modification, particularly that in *Noun Complement* construction research.

Secondly, this study also develops a comprehensive picture of the linguistic and extralinguistic aspects of how the target structure functions differently in the research articles of writers from cross-disciplinary and cross-cultural backgrounds. Specifically, in terms of the linguistic aspect, these two groups of writers differ statistically in applying the functions of the target structure to introduce or signal identity relationships, elaborative interpretations, complex grammatical structures, dependent *wh*-clauses, and substantive expressions, as the qualifiers of the stance noun and concurrently the subjects of its complement, transition markers, frame markers, endophoric markers, code glosses, or for applying the information principle and the principle of end-weight. As to the extra-linguistic functions, they diverge statistically on deploying the target structure to serve as hedges, boosters, attitude markers, to attribute stance nouns to abstract entities, other humans, or writers themselves, to display informality and launch strategic effort, or even to both overtly aver the stance noun and refer it to other parts of the text.

Finally, the present study reveals the underlying motivations and enabling factors behind the different use of post-modification in the *Noun Complement* construction by soft and hard field writers from international and local academic journals. Specifically, the reasons that underlie the cross-cultural variations have been attributed to the following seven major factors in four dimensions: namely, (a) idiosyncratic features of the semantic,

DOI:10.6814/NCCU202100169

syntactic and punctuation systems in the Chinese language (e.g., interchangeable use of dashes and colons), (b) Chinese EFL writers' language competence in English academic writing (i.e., low levels) in the linguistic dimension, (c) traditions of Chinese rhetoric (e.g., emphasis on the use of "fancy syntax"), (d) Chinese EFL writers' academic competence in English academic writing (e.g., rhetorical unawareness of balancing the use of hedges and boosters) in the rhetorical dimension, (e) notions of writer-responsibility in the West versus reader-responsibility in the East in the relational dimension (e.g., ignoring the guidance of readers through text) in the relational dimension, (f) Chinese cultural practices rooted in Confucian and Taoist traditions (e.g., no need for published writers to validate themselves as knowledgeable others), and (g) influence of social and political movements in Western cultures (e.g., less impacts of activism) in the socio-cultural dimension.

In the similar vein, the motivations for the cross-disciplinary differences in the use of the target structure were explained based on the contributing factors in the same four dimensions as follows: (a) linguistic conventions in the genre of research articles (e.g., the extent to which nominalization is adhered to), (b) nature of the grammatical complexity differences (e.g., hard scientists more favoring phrasal complexity) in the linguistic dimension, (c) modes of knowledge construction (e.g., knowledge constructed based on reasoning and persuasion in soft fields), (d) disciplinary rhetorical traditions (e.g., less hedging in the natural sciences) in the rhetorical dimension, (e) interpersonal writer-reader relationship (e.g., social science writers' preference for reader-oriented functions) in the relational dimension, (f) developmental differences in scientific research (e.g., more growing interdisciplinary research trend in hard domains), and (g) influence of social and political movements in Western cultures (e.g., post-modern and feminist movements more deeply affecting social scientists) in the social-cultural dimension. As such, the explanatory model built on the linguistic, rhetorical, relational, and socio-cultural aspects of how the target structure works across disciplinary fields and cultures in this study, in fact, situates any discipline or culture in a wider social world and acknowledges the concerted effort of the lingua-rhetorical and cultural-pragmatic factors to mould the variations between writers

DOI:10.6814/NCCU202100169

from different disciplines and cultures. Based on this, the reasons identified in this research thus extend far beyond that of the disciplines and ESL/EFL writers per se in prior studies.

6.3 Implications of the present study

For researchers interested in the differences in disciplinary and cultural use of stance constructions, this study has demonstrated the wide use of a grammatically marked structure that syntactically and semantically breaks apart head noun and its complement clause – namely, post-modification in *Noun Complement* construction – across disciplinary fields and writers of distinct linguistic-cultural backgrounds. The research value of the target structure has been proved by the cross-disciplinary and cross-cultural variations in its rich and varied grammatical forms, its capability for playing a wide variety of linguistic and extra-linguistic roles, and the underlying motivations behand its use. As such, this study contributes to the theory of stance construction in academic writing, by proposing that it is worth looking at any grammatical component in a stance-making structure in terms of its forms, functions, and the reasons behind its use to explore its research value.

In terms of methodological implications, the present study has illustrated the feasibility of applying the inferential statistical methods in *Noun Complement* construction research to identify the variances in the lexico-grammatical features of the target structure between disciplinary fields and writers of distinct linguistic-cultural backgrounds. Similarly, the use of the optimized combination of regular expression queries in this study to efficiently and accurately extract the cases of the target structure can be useful for the corpus-based research with the need to extract a large number of grammatically complex structures, such as the post-modification in the *Noun Complement* construction in the present study.

The linguistic findings of the present study may be especially helpful to instruction in English for academic purposes (EAP) writing classes. The sets of lexico-grammatical features identified in the use of post-modification in *Noun Complement* construction can be explicitly taught and excised, whereby student writers can relate these features to the corresponding functions they perform. Otherwise, the relatively hidden grammatical position of the target structure (i.e., between head noun and its complement clause) is likely

DOI:10.6814/NCCU202100169

to escape the notice of the students as novice research writers, and thus cannot be acquired by them through implicit learning. The same is true for the explicit instruction on the underlying motivations for choosing such features to serve relevant functions. Specifically, the reasons behind cross-disciplinary differences can be shared with the students from different disciplines, while that underlying cross-cultural differences can likewise be done for the students with different first languages and cultural backgrounds (e.g., EFL student writers in China). Of course, EFL student writers from various departments at the university can be concurrently initiated into all of these reasons, taking heed to cross-disciplinary or cross-cultural differences or both as needed or as desired. Moreover, the lexico-grammatical features and functions of the target structure, along with the reasons that underlie them, relate correspondingly to its linguistic, rhetorical, relational, and socio-cultural aspects. Hence, sentences or paragraphs like those in the research articles analyzed in this dissertation can be used, either with a linguistic focus to illustrate the textual functions of particular lexico-grammatical features or, with a less linguistic focus, to explain the rhetorical efforts, relational concerns, and socio-cultural implications beyond the surface lexico-grammatical level. It should be also noted here that certain functions are often less than transparent and may be challenging for students, in particular EFL students, to comprehend, and thus require special instruction. To illustrate, the function of *qualifier + subject* performed by the lexico-grammatical feature of *prepositional phrase* entails the prepositional phrase serving simultaneously as the qualifier of its preceding stance noun and as the subject of its following proposition (as *to NASA* in *by the 1998 Congressional directive to NASA to protect the Earth and its inhabitants from the threat of asteroid impact*). These particular functions may need to be analyzed directly by students until the relationships involved become abundantly clear for students to draw on them effectively.

Secondly, language teachers or EAP instructors would need to set an example of not taking the stereotypical view about the use of any grammatical structure based on unreliable native-speaker or ESL/EFL-learner intuitions (given part of language teachers or EAP instructors as ESL/EFL learners themselves). In the case of the present study,

DOI:10.6814/NCCU202100169

although one would intuitively expect post-modification in *Noun Complement* construction to be seldom used, it actually maintains wide popularity in the stance repertoire of academic writers in various disciplinary domains and with different first languages and cultural backgrounds. Similarly, they may also need to withdraw a related intuitive stereotype-based judgment as to the most used lexico-grammatical features of post-modification use. In fact, even though a grammatical form might be expected by intuition to be most used in the target structure, other uses are in fact equally or more common. More importantly, language teachers or EAP instructors are supposed to possess a comprehensive understanding of the cross-disciplinary and cross-cultural differences in English academic writing themselves, in order to provide systematic and intentional writing instruction. Since they are teaching professionals with highly advanced English proficiency, there should be no difficulty for them in gaining a good command of the lexico-grammatical features of the target structure and their corresponding linguistic, rhetorical, relational, and socio-cultural functions. However, the major obstacle in the way of achieving successful and effective instruction is whether they are able to adequately grasp the underlying reasons beyond the use of these features by writers from soft and hard fields and with contrasting first languages and cultural backgrounds. In terms of cross-disciplinary differences, they may need to obtain an in-depth understanding of the modes of knowledge construction, the rhetorical traditions, the Anglo-Saxon norms of scientific writing, the preferences between phrasal and clausal complexity, the complexity of the current research, the influence of social and political movements, in the fields of both natural and social sciences. As to the cross-cultural differences, the focus is on becoming sufficiently familiar with the features of the semantic, syntactic and punctuation systems, the rhetorical traditions, the aesthetic features, and the cultural practices of learners' first languages, along with their specific writing needs (e.g., changing the stylistic habit of overusing flowery language and big words). This may seem to be a huge amount of work for language teachers or EAP instructors. Nevertheless, this dissertation provides a comprehensive picture of cross-disciplinary and cross-cultural differences in how post-modification in *Noun Complement*

DOI:10.6814/NCCU202100169

construction functions with systematic explanations from the linguistic, rhetorical, relational, and socio-cultural aspects, which can, in turn, aid these teaching professionals in easing the workload. For example, the sentences or paragraphs of the research articles analyzed in the present study can be directly used in their classroom teaching.

Thirdly, the application of corpus concordance tools could be regarded as a highly effective way to initiate students into the use of post-modification in *Noun Complement* construction, given that "corpora and concordance packages present very useful resources for the creation of exercises that motivate the learner and promote her/his language awareness" (Römer, 2009: p. 92). Specifically, learners can be instructed to search the target structure in research articles from varying subject areas and by writers with different linguistic and cultural backgrounds. This can be done in the existing authoritative online corpora (e.g., Corpus of Contemporary American English (COCA) or the British National Corpus (BNC)), where academic genres and authorship are usually selectable as subcorpora with a list of disciplines and authorial characteristics, or in self-built corpora by instructors where applicable, through the built-in concordance tools or the corpus software *AntConc* with research articles converted into plain text format. Follow-up activities include but are not limited to identifying the types of lexico-grammatical features of the target structure and the correlation between such features and their linguistic, rhetorical, relational, and socio-cultural functions within the discourse context of the entire research article (concordance tools usually allow users to look at the searching item in its context). Hand-on experiences of this kind are likely to provide learners with access to authentic language materials and to raise their language awareness through better noticing. It is worth noting here that the target students are supposed to be EAP, ESP, or advanced ESL learners, because these activities entail a good master of analytical skills, thereby requiring an adequate command of English.

Lastly, this dissertation is the first to systematically examine the use of post-modification in *Noun Complement* construction by published academic writers of English-medium journal articles. Thus, it is well suited for incorporating into the grammar books

DOI:10.6814/NCCU202100169

that cover information on cross-disciplinary or cross-cultural variation and sensitivity of various linguistic constructions, such as Biber et al.'s *Longman Grammar of Spoken and Written English* (1999). In so doing, these books can provide a fuller picture of the uses, forms, and functions of *Noun Complement* construction across different disciplinary and cultural contexts, thus complementing the systematic accounts of nominal modification use. More importantly, they will help learners to gain a better understanding of the actual and more complete use of nominal modification in specific occasions of writing.

6.4 Limitation and suggestions for future research

As with any study, the present study has its limitations. The limitations and findings of this dissertation can suggest a number of areas for future studies. In continuation of the component of this study focusing on the corpus-based and inferential statistical approach to characterizing the lexico-grammatical features of post-modification use in the *Noun Complement* construction of published academic writing, further differentiation and description of such features of the target structure is the first accessible goal. In particular, full morphosyntactic descriptions of less frequently used features (e.g., *extraposition*) will not only define this pattern of use more accurately, but also pave the way for conducting more precise identification and grouping of lexico-grammatical features in future corpus-based research.

Moreover, the vast majority of linguistic, rhetorical, relational, and socio-cultural functions served by these lexico-grammatical features were grouped based on either the cross-disciplinary or the cross-cultural differences in the present study. While it indeed may be possible that these groupings do simply represent the most common functional accounts available in published research articles from varying disciplinary fields and by writers with contrasting linguistic and cultural backgrounds, it seems worthwhile to draw finer-grained distinctions within these groupings (e.g., subdividing the function category of *abstract entity* into a number of further groups of *data*, *research methods*, *institutions*, etc.) in related follow-up research.

DOI:10.6814/NCCU202100169

This dissertation reveals the underlying reasons beyond the cross-disciplinary and cross-cultural differences from linguistic, rhetorical, relational, and socio-cultural perspectives. It appears to be an adequate and systematic analytical tool for analyzing the reasons for different uses of post-modification in *Noun Complement* construction in research articles. However, although this study prioritized collecting a sufficiently large sample of academic texts (600 files, amounting to 5,567,545 words), future researchers will, through the construction of a larger corpus, be able to find out additional reasons to improve, develop and reorganize the present explanatory model.

While the corpus explored here represents the writing of published expert writers from soft and hard research fields and with contrasting first languages and cultural backgrounds, it may be effective to manipulate a corpus to characterize the writing of ESL/EFL learners at different levels (e.g., academic texts of student writers from undergraduate to doctoral levels or at different levels
of proficiency in English) and from varying university departments, and to compare it with the corpus built for this dissertation. Using the categories of lexico-grammatical features, functions, and the analytical tool for the analysis of underlying reasons, these future studies are suggested not only to confirm the findings of this study as disciplinary and cultural features, but also to examine the degree to which the explanatory model of lexico-grammatical features, functions, and underlying reasons developed in this dissertation are stable against the representation of the writing from the authors with more complicated academic, linguistic and cultural backgrounds.

Finally, much of the corpus and mixed-methods analysis here, both methods and findings, is likely to be helpful for replicated or refined studies carrying these methods over to other academic genres (e.g., conference papers, book reviews, laboratory research reports, textbooks, etc.). Later researchers could target particular genres and provide an even fuller picture of variations in lexico-grammatical features, functions, and underlying reasons beyond the genre of research articles. Similarly, these ensuing studies are expected

DOI:10.6814/NCCU202100169

to affirm the findings here as genre features, while also investigating the stability of the present explanatory model across other genres.

DOI:10.6814/NCCU202100169

REFERENCES

Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. Amsterdam: Benjamins.

Adobe Inc. (2020). *Adobe Acrobat Pro DC (Version 20.006.20034)* [Computer Software]. San Jose, CA: Adobe Inc.

Akin, H., & Scheufele, D. A. (2017). Overview of the science of science communication. In K. H. Jamieson, D. Kahan, & D. A. Scheufele (Eds.), *The Oxford Handbook of the Science of Science Communication* (pp. 25-33). New York: Oxford University press.

Aktas, R. N., & Cortes, V. (2008). Shell nouns as cohesive devices in published and ESL student writing. *Journal of English for Academic Purposes*, *7*(1), 3-14.

Allison, D. (1995). Assertions and alternatives: Helping ESL undergraduates extend their choices in academic writing. *Journal of Second Language Writing*, *4*(1), 1-15.

Anthony, L. (2015). *TagAnt (Version 1.2.0)* [Computer Software]. Tokyo: Waseda University.

Anthony, L. (2019). *AntConc (Version 3.5.8)* [Computer Software]. Tokyo: Waseda University.

Ariana, S. M. (2010). Some thoughts on writing skills. *Annals of the University of Oradea, Economic Science Series*, *19*(1), 134-140.

Ashenden, S. (1997). Feminism, postmodernism and the sociology of gender. In D. Owen (Ed.), *Sociology after Postmodernism* (pp. 40-64). London: SAGE.

Aull, L. L. (2015). *First-Year University Writing: A Corpus-Based Study with Implications for Pedagogy*. New York: Palgrave Macmillan.

Bacon, N. (2013). Style in academic writing. In M. Duncan & S. Vanguri (Eds.), *The Centrality of Style* (pp. 173-190). Anderson, SC: Parlor Press.

Bazerman, C. (1988). *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison, WI: University of Wisconsin Press Madison.

Bazerman, C. (1993). Foreword, in N. Blyler and C. Thralls (Eds.), *Professional Communication: The Social Perspective* (pp. vii-x). Newbury Park, CA: Sage.

DOI:10.6814/NCCU202100169

Becher, T. (1989). *Academic Tribes and Territories: Intellectual Enquiry and the Culture of Disciplines.* Milton Keynes: SHRE/Open University Press.

Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher Education, 19*(*2*), 151-161.

Becher, T., & Trowler, P. (2001). *Academic Tribes and Territories: Intellectual Enquiry and the Culture of Disciplines*. Buckingham: Society for Research into Higher Education & Open University Press.

Benson, C. (2002). Transfer/Cross-linguistic influence. *ELT Journal*, *56*(1), 68-70.

Biber, D. (1991). *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, D. (2006a). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, *5*(2), 97-116.

Biber, D. (2006b). *University Language: A Corpus-based Study of Spoken and Written Registers*, vol. 23. Amsterdam: John Benjamins.

Biber, D., & Clark, V. (2002). Historical shifts in modification patterns with complex noun phrase structures. In T. Fanego, M. J. Lépez-Couso, & J. Perez-Guerra (Eds.), *English Historical Morphology* (pp. 43-66). Amsterdam: John Benjamins.

Biber, D., Connor, U. & Upton, T. (2007). *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*. Amsterdam: John Benjamins.

Biber, D., & Conrad, S. (1999). Lexical Bundles in Conversations and Academic Prose. In H. Hasselgård & S. Oksefjell (Eds.), *Out of Corpora: Studies in Honour of Stig Johansson* (pp. 181-190). Amsterdam: Rodopi.

Biber, D., & Conrad, S. (2014). *Variation in English: Multi-Dimensional Studies*. London: Routledge.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at …: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371-405.

Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A multidimensional comparison. *TESOL Quarterly*, *36*(1), 9-48.

DOI:10.6814/NCCU202100169

Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text*, *9*(1), 93-124.

Biber, D., & Gray, B. (2013). Nominalizing the verb phrase in academic science writing. In B. Aarts, J. Close, G. Leech, & S. Wallis (Eds.), *The Verb Phrase in English: Investigating Recent Language Change with Corpora* (pp. 99-132). Cambridge: Cambridge University Press.

Biber, D., & Gray, B. (2016). *Grammatical Complexity in Academic English: Linguistic Change in Writing*. Cambridge: Cambridge University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, *45*(1), 5-35.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Written and Spoken English*. Harlow: Longman.

Bizzell, P. (1992). *Academic Discourse and Critical Consciousness*. Pittsburgh: University of Pittsburgh Press.

Bloch, J. and Chi, L. (1995). A comparison of the use of citations in Chinese and English academic discourse. In D. Belcher & G. Braine (Eds.), *Academic Writing in a Second Language: Essays on Research and Pedagogy* (pp. 213-274). Norwood, NJ: Ablex Publishing Company.

Bodde, B., (1991). *Chinese Thought, Society and Science: The Intellectual and Social Background of Science and Technology in Pre-modern China*. Honolulu: University of Hawaii Press.

Bolton, K., Nelson, G., & Hung, J. (2002). A corpus-based study of connectors in student writing: Research from the International Corpus of English in Hong Kong (ICE-HK). *International Journal of Corpus Linguistics*, *7*(2), 165-182.

Brown, P. & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

DOI:10.6814/NCCU202100169

Bühler, K. (1990 [1934]). *Theory of Language: The Representational Function of Language*. Translation of *Sprachtheorie* by Donald Fraser Goodwin. Amsterdam: John Benjamins.

Byrnes, H. (2009). Emergent L2 German writing ability in a curricular context: A longitudinal study of grammatical metaphor. *Linguistics and Education, 20*, 50-66.

Cai, G. (1999). Texts in contexts: Understanding Chinese students' English compositions. In C. R. Cooper & L. Odell (Eds.), *Evaluating Writing: The Role of Teachers' Knowledge about Text, Learning, and Culture* (pp. 279-297). Urbana, IL: National Council of Teachers of English.

Cameron, D. (2012). Epilogue. In K. Hyland & C. S. Guinda (Eds.), *Stance and Voice in Written Academic Genres* (pp. 249-256). London: Palgrave Macmillan.

Chafe, W. (1986). Evidentially in English conversation and academic writing. In W. Chafe & J. Nichols (Eds.), *Evidentiality: The Linguistic Coding of Epistemology* (pp. 261-272). Norwood, NJ: Ablex.

Chafe, W. & Nichols J. (1986). *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex Publishing Corporation.

Channell, J. (1990). Precise and vague expressions in writing on economics. In W. Nash (Ed.), *The Writing Scholar: Studies in Academic Discourse* (pp. 95-117). Newbury Park, CA: Sage.

Charles, M. (2003). A corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes*, *2*(4), 313-326.

Charles, M. (2006). The construction of stance in reporting clauses: A cross-disciplinary study of theses. *Applied Linguistics*, *27*(3), 492-518.

Charles, M. (2007). Argument or evidence? Disciplinary variation in the use of the noun "that" pattern in stance construction. *English for Specific Purposes*, *26*(2), 203-218.

DOI:10.6814/NCCU202100169

Chen, Q. S. (2014). Exploring the literacy practices of undergraduate English majors' academic writing in China. In *Paper presented at the postgraduate research conference 2013-2014, Hong Kong*.

Chen, R. (2020). Single author self-reference: Identity construction and pragmatic competence. *Journal of English for Academic Purposes, 45*, Article 100856.

Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14*(2), 30-49.

Clarke, B. L. (1964). Multiple Authorship Trends in Scientific Papers. *Science, 143*(3608), 822-824.

CNKI-EJCR. (2017). *Report on the International & Domestic Citation of China's English Academic Journals*.

Coffin, C. (2004). Arguing about how the world is or how the world should be: The role of argument in IELTS tests. *Journal of English for Academic Purposes, 3*(3), 229-246.

Cohen, J. (1977). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Academic Press, INC.

Conrad, S. (1996). Academic discourse in two disciplines: Professional writing and student development in Biology and History. Unpublished PhD thesis, Northern Arizona University.

Cortes, V. (2002). Lexical bundles in Freshman composition. In R. Reppen, S. M. Fitzmaurice, & D. Biber (Eds.), *Using Corpora to Explore Linguistic Variation* (pp. 131-145). Amsterdam: John Benjamins Publishing Company.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23*(4), 397-423.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.

Crismore, A., Markkanen, R., & Steffensen, M. S. (1993). Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students. *Written communication, 10*(1), 39-71.

DOI:10.6814/NCCU202100169

Crompton, P. (1997). Hedging in academic writing: Some theoretical problems. *English for Specific Purposes*, *16*(4), 271-287.

Crosthwaite, P., Cheung, L., & Jiang, F. K. (2017). Writing with attitude: Stance expression in learner and professional dentistry research reports. *English for Specific Purposes*, *46*, 107-123.

Culicover, P. W., & Rochemont, M. S. (1990). Extraposition and the complement principle. *Linguistic Inquiry*, *21*(1), 23-47.

Curran, T. D. (2014). A response to Professor Wu Zongjie's 'Interpretation, autonomy, and transformation: Chinese pedagogic discourse in a cross-cultural perspective'. *Journal of Curriculum Studies*, *46*(3), 305-312.

Davis-Floyd, R., & Sargent, C. F. (1997). *Childbirth and Authoritative Knowledge: Cross-Cultural Perspectives*. Berkeley: University of California Press.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*(4), 499-533.

Demir, C. (2019). The needless complexity in academic writing: Simplicity vs. flowery. *The Reading Matrix: An International Online Journal*, *19*(2), 13-27.

Ding, Y. & Miao, W. (2020). The status quo, development and challenges of Chinese English academic journals going global. *Publication Research*, *6*, 65-69.

Dressen, D. (2003). Geologists' implicit persuasive strategies and the construction of evaluative evidence. *Journal of English for Academic Purposes*, *2*(4), 273-290.

Du, J. (2020). *Non-Native English-Speaking Engineers' Writing at the Workplace*. Singapore: Springer.

Egbert, J. (2014). Reader perceptions of linguistic variation in published academic writing. Unpublished PhD thesis, Northern Arizona University.

Egbert, J. (2015). Publication type and discipline variation in published academic writing: Investigating statistical interaction in corpus data. *International Journal of Corpus Linguistics*, *20*(1), 1-29.

DOI:10.6814/NCCU202100169

Ellis, N. C. (2008). The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *The Modern Language Journal*, *92*(2), 232-249.

Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Elman, B. A. (2000). *A Cultural History of Civil Examinations in Late Imperial China*. Berkeley, CA: University of California Press.

Enfield, N. J. (2009). Language: Social motives for syntax. *Science*, *324*(*5923*), 39.

Ernst, D. (2005). The complexity and internationalization of innovation: The root causes. In UNCTAD (Ed.), *Globalization of R&D and Developing Countries: Proceedings of the Expert Meeting* (pp. 61-87). Geneva: UNCTAD.

Fahnestock, J. (1986). Accommodating science: The rhetorical life of scientific facts. *Written Communication*, *3*(3), 275-296.

Fanelli, D. (2020). Pressures to publish: What effects do we see? In M. Biagioli & A. Lippman (Eds.), *Gaming the Metrics: Misconduct and Manipulation in Academic Research* (pp. 111-122). Cambridge, MA: MIT Press.

Feng, Y., & Zhou, R. (2007). Using hedges in academic writings: A comparative study. *Foreign Language and Literature Studies*, *24*(2), 108-112.

Field, A. (2009). *Discovering Statistics Using SPSS*. London: Sage.

Fløttum, K., Dahl, T., & Kinn, T. (2006). *Academic Voices: Across Languages and Disciplines*. Amsterdam: John Benjamins.

Flowerdew, J. (2003). Signalling nouns in discourse. *English for Specific Purposes*, *22*(4), 329-346.

Flowerdew, J., & Li, Y. (2009). English or Chinese? The trade-off between local and international publication among Chinese academics in the humanities and social sciences. *Journal of Second Language Writing*, *18*(1), 1-16.

Francis, E. J. (2010). Grammatical weight and relative clause extraposition in English. *Cognitive Linguistics*, *21*(1), 35-74.

DOI:10.6814/NCCU202100169

Francis, G. (1986). *Anaphoric Nouns*. Birmingham: English Language Research, University of Birmingham.

Francis, G. (1994). Labelling discourse: an aspect of nominal-group lexical cohesion. In M. Coulthard (Ed.), *Advances in Written Text Analysis* (pp. 83-101). London: Routledge.

Gao, X. (2016). A cross-disciplinary corpus-based study on English and Chinese native speakers' use of linking adverbials in academic writing. *Journal of English for Academic Purposes*, *24*, 14-28.

Gao, Y., & Wen, Q. (2009). Co-responsibility in the dialogical co-construction of academic discourse. *TESOL Quarterly*, *43*(4), 700-703.

Gardezi, S. A., & Nesi, H. (2009). Variations in the writing of economics students in Britain and Pakistan: The case of conjunctive ties. In M. Charles, D. Pecorari, & S. Hunston (Eds.), *Academic Writing at the Interface of Corpus and Discourse* (pp. 236-250). London: Continuum.

Garry, A., & Pearsall, M. (2015). *Women, Knowledge, and Reality: Explorations in Feminist Philosophy*. London: Routledge.

Gilbert, G. (1976). The transformation of research findings into scientific knowledge. *Social Studies of Science, 6*(3-4), 281-306.

Gilbert, G. N., & Mulkay, M. (1984). *Opening Pandora's Box: A Sociological Analysis of Scientists' Discourse*. Cambridge: Cambridge University Press.

Goffman, E. (1967). *Interaction Ritual: Essays on Face-to-Face Behavior*. New York: Pantheon.

Gray, B. (2011). Exploring academic writing through corpus linguistics: When discipline tells only part of the story. Unpublished PhD thesis, Northern Arizona University.

Gray, B. (2015). *Linguistic Variation in Research Articles: When Discipline Tells Only Part of the Story*. Amsterdam: John Benjamins.

DOI:10.6814/NCCU202100169

Gross, A. G., & Chesley, P. (2012). Hedging, stance and voice in medical research articles. In K. Hyland & C. S. Guinda (Eds.), *Stance and Voice in Written Academic Genres* (pp. 85-100). London: Palgrave Macmillan.

Guinda, C. S. & Hyland, K. (2012). Introduction: A context-sensitive approach to stance and voice. In K. Hyland & C. S. Guinda (Eds.), *Stance and Voice in Written Academic Genres* (pp. 1-11). London: Palgrave Macmillan.

Hall, E.T. (1976). *Beyond Culture*. New York: Doubleday.

Halliday, M. A. K. (1994). Spoken and written modes of meaning. In D. Graddol & O. Boyd-Barrett (Ed.), *Media Texts: Authors and Readers: A Reader* (pp. 51-73). Clevedon: Multilingual Matters.

Halliday, M. A. K. (1973). *Explorations in the Functions of Language*. London: Edward Arnold.

Halliday, M. A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meanings*. London: Edward Arnold.

Halliday, M. A. K. (1979). Differences between spoken and written language: Some implications for language teaching. In G. Page et al. (Eds.), *Communication through Reading: Proceedings of the 4th Australian Reading Conference* (pp. 37-52). Adelaide: Australian Reading Association.

Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.

Halliday, M. A. K. (1998). Things and relations: Regrammaticising experience as technical knowledge. In J. R. Martin & R. Veel (Eds.), *Reading Science: Critical and Functional Perspectives on Discourses of Science* (pp. 185-235). London: Routledge.

Halliday, M. A. K. (2013). Meaning as choice. In L. Fontaine, T. Bartlett, & G. O'Grady (Eds.), *Systemic Functional Linguistics: Exploring Choice*. Cambridge: Cambridge University Press.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

DOI:10.6814/NCCU202100169

Halliday, M. A. K., & Martin, J. R. (1993/1996). *Writing Science, Literacy and Discursive Power*. Pittsburgh, PA: University of Pittsburgh Press.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (1999). *Construing Experience Through Meaning: A Language-Based Approach to Cognition*. London: Cassell.

Hasselgård, H., Johansson, S., & Lysvåg, P. (1998). *English Grammar: Theory and Use*. Oslo: Universitetsforlaget.

Hewings, A. (2012). Stance and voice in academic discourse across channels. In K. Hyland & C. S. Guinda (Eds.), *Stance and Voice in Written Academic Genres* (pp. 187-201). London: Palgrave Macmillan.

Hinds, J. (1987). Reader versus writer responsibility: A new typology. In U. Connor & R. B. Kaplan (Eds.), *Writing Across Languages: Analysis of L2 Text* (pp. 141-152). Reading, Massachusetts: Addison Wesley.

Hinds, J. (2001). Reader versus writer responsibility: A new typology. *Landmark Essays on ESL Writing*, *17*, 63-74.

Hinkel, E. (1994). Native and nonnative speakers' pragmatic interpretations of English texts. *TESOL Quarterly*, *28*(2), 353-376.

Hinkel, E. (1997). Indirectness in L1 and L2 academic writing. *Journal of Pragmatics*, *27*(3), 361-386.

Hinkel, E. (1999). Objectivity and credibility in L1 and L2 academic writing. In E. Hinkel & M. H. Long (Eds.), *Culture in Second Language Teaching and Learning* (pp. 90-108). Cambridge: Cambridge University Press.

Hinkel, E. (2005). Hedging, inflating, and persuading in L2 academic writing. *Applied Language Learning*, *15*(1-2), 29-53.

Hockin, J., Miller, C., & Magee, P. T. (1996). *Writing as a Way of Knowing in a Cross-Disciplinary Classroom*. Minneapolis, MN: Center for Interdisciplinary Studies of Writing, University of Minnesota.

Holmes, J. (1984). Modifying illocutionary force. *Journal of Pragmatics, 8*(3), 345-365.

Holmes, J. (1988). Doubt and certainty in ESL textbooks. *Applied Linguistics, 9*(1), 21-44.

DOI:10.6814/NCCU202100169

Holmes, J. (1995). *Women, Men and Politeness.* London: Longman.

Hu, G., & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English-and Chinese-medium journals. *Journal of Pragmatics*, *43*(11), 2795-2809.

Hu, Y. Y. (2005). Western and Chinese thinking differences and English learning. *Digest of Management Science*, *7*, 53-54.

Huang, R. (2007). Semantic prosody of adjective amplifiers: An analysis of Chinese learners of English. *Foreign Language Education, 28*(4), 57-60.

Huddleston, R. D., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hunston, S. (1989). Evaluation in experimental research articles. Unpublished PhD thesis, University of Birmingham.

Hunston, S. (1993). Evaluation and ideology in scientific writing. In M. Ghadessy (Ed.), *Register Analysis: Theory and Practice* (pp. 57-73). London: Pinter.

Hunston, S. (1994). Evaluation and organization in a sample of written academic discourse. In M. Coulthard (Ed.), *Advances in Written Text Analysis* (pp. 191-218). London: Routledge.

Hunston, S., & Francis, G. (1999). *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins.

Hutauruk, B. S. (2017). Errors of punctuation transfer in English compositions by English learners Academic Year (2013) English Department in FKIP UHN Pematangsiantar. *Journal of English Language and Culture*, *6*(1), 1-18.

Hyland, K. (1996). Writing without conviction? Hedging in science research articles. *Applied linguistics*, *17*(4), 433-454.

Hyland, K. (1997). Language attitudes at the handover: Communication and identity in 1997 Hong Kong. *English World-Wide*, *18*(2), 191-210.

Hyland, K. (I998a). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics, 30*(4), 437-455.

DOI:10.6814/NCCU202100169

Hyland, K. (1998b). Exploring corporate rhetoric: Metadiscourse in the CEO's letter. *Journal of Business Communication, 35*(2), 224-245.

Hyland, K. (1998c). *Hedging in Scientific Research Articles*. Amsterdam: John Benjamins Publishing.

Hyland, K. (1999a). Disciplinary discourses: Writer stance in research articles. In C. Candlin & K. Hyland (Eds.), *Writing: Texts, Processes and Practices* (pp. 99-121). London: Longman.

Hyland, K. (1999b). Talking to students: Metadiscourse in introductory textbooks. *English for Specific Purposes, 18*(1), 3-26.

Hyland, K. (2000). Hedges, boosters and lexical invisibility: Noticing modifiers in academic texts. *Language Awareness, 9*(4), 179-197.

Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes, 20*(3), 207-226.

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics, 34*(8), 1091-1112.

Hyland, K. (2003). *Second Language Writing*. Cambridge: Cambridge University Press.

Hyland, K. (2004). *Disciplinary Discourses: Social Interactions in Academic Writing*. Ann Arbor, MI: University of Michigan Press.

Hyland, K. (2005a). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173-192.

Hyland, K. (2005b). *Metadiscourse: Exploring Interaction in Writing*. London: Continuum.

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing, 16*(3), 148-164.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27*(1), 4-21.

Hyland, K. (2009). *Academic Discourse: English in a Global Context*. London: Continuum.

Hyland, K. (2012). *Disciplinary Identities: Individuality and Community in Academic Discourse*. Cambridge: Cambridge University Press.

DOI:10.6814/NCCU202100169

Hyland, K. (2013). Innovating instruction: English in the discipline at the University of Hong Kong. *Hong Kong Journal of Applied Linguistics*, *14*(2), 3-19.

Hyland, K. (2014). English for academic purposes. In C. Leung & B. V. Street (Eds.), *The Routledge Companion to English Studies* (pp. 392-404). London: Routledge.

Hyland, K. (2015). *Teaching and Researching Writing*. London: Routledge.

Hyland, K. (2019). *Metadiscourse: Exploring Interaction in Writing*. London: Bloomsbury Academic.

Hyland, K., & Guinda, C. S. (2012). *Stance and Voice in Written Academic Genres*. Hampshire: Palgrave Macmillan.

Hyland, K., & Jiang, K. F. (2016). Change of attitude? A diachronic study of stance. *Written Communication*, *33*(3), 251-274.

Hyland, K., & Jiang, K. F. (2018). "In this paper we suggest": Changing patterns of disciplinary metadiscourse. *English for Specific Purposes*, *51*, 18-30.

Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, *6*(2), 183-206.

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied linguistics*, *25*(2), 156-177.

Hyland, K., & Tse, P. (2005a). Evaluative *that* constructions: Signalling stance in research abstracts. *Functions of Language*, *12*(1), 39-64.

Hyland, K., & Tse, P. (2005b). Hooking the reader: A corpus study of evaluative *that* in abstracts. *English for Specific Purposes*, *24*(2), 123-139.

IBM Corp. (2018). *IBM SPSS Statistics for Windows (Version 24.0)* [Computer Software]. Armonk, NY: IBM Corp.

Ivanič, R. (1991). Nouns in search of a context: A study of nouns with both open-and closed-system characteristics. *International Review of Applied Linguistics in Language Teaching*, *29*(2), 93-114.

Jiang, K. F. (2015). Nominal stance construction in L1 and L2 students' writing. *Journal of English for Academic Purposes*, *20*, 90-102.

DOI:10.6814/NCCU202100169

Jiang, K. F. (2017). Stance and voice in academic writing. *International Journal of Corpus Linguistics*, *22*(1), 85-106.

Jiang, K. F., & Hyland, K. (2015). 'The fact that': Stance nouns in disciplinary writing. *Discourse Studies*, *17*(5), 529-550.

Kaplan, R. (1966). Cultural thought patterns in intercultural education. Language Learning, *16*(1), 1-20.

Kim, C., & Crosthwaite, P. (2019). Disciplinary differences in the use of evaluative that: Expression of stance via that-clauses in business and medicine. *Journal of English for Specific Purposes*, *41,* 1-14.

Kim, L. C., & Lim, J. M. H. (2013). Metadiscourse in English and Chinese research article introductions. *Discourse Studies*, *15*(2), 129-146.

Kirkpatrick, A., & Mulligan, D. (2002). Cultures of learning: Critical reading in the social and applied sciences. *Australian Review of Applied Linguistics*, *25*(2), 73-99.

Knorr-Cetina, K. (1981). *The Manufacture of Knowledge.* Oxford: Pergamon Press.

Kress, G. (1989). *Linguistic Processes in Sociocultural Practice.* Oxford: Oxford University Press.

Kwiek, M. (2012). The growing complexity of the academic enterprise in Europe: A panoramic view. *European Journal of Higher Education*, *2*(2-3), 112-131.

Lancaster, Z. (2016). Expressing stance in undergraduate writing: Discipline-specific and general qualities. *Journal of English for Academic Purposes*, *23*, 16-30.

Latour, B. (1987). *Science in Action*. Cambridge, MA: Harvard University Press.

Lee, W. O., & Tan, J. P. (2018). The new roles for twenty-first-century teachers: Facilitator, knowledge broker, and pedagogical weaver. In H. Niemi, A. Toom, A. Kallioniemi, & J. Lavonen (Eds), *The Teacher's Role in the Changing Globalizing World: Resources and Challenges Related to the Professional Work of Teaching* (pp. 11-31). Leiden: Brill Sense.

Leech, G. (1981, 1995). *Semantics: The Study of Meaning*. Harmondsworth: Penguin.

DOI:10.6814/NCCU202100169

Lei, L. (2012). Linking adverbials in academic writing on applied linguistics by Chinese doctoral students. *Journal of English for Academic Purposes, 11*(3), 267-275.

Li, F. (2018). For explaining or summarizing: Differences between colon and dash. *Chinese Monthly, 4*, 93.

Lillis, T. M. (2012). Economies of signs in writing for academic publication: The case of English medium "national" journals. *Journal of Advanced Composition, 32*(3-4), 695-722.

Lillis, T. M., & Curry, M. J. (2010). *Academic Writing in a Global Context: The Politics and Practices of Publishing in English*. London: Routledge.

Lim, J. M. H. (2017). Writing descriptions of experimental procedures in language education: Implications for the teaching of English for academic purposes. *English for Specific Purposes, 47*, 61-80.

Lin, H. (2000). *Learning and Application of Punctuation.* Beijing: People's Publishing House.

Liu, Y., & Du, Q. (2018). Intercultural rhetoric through a learner lens: American students' perceptions of evidence use in Chinese yìlùnwén writing. *Journal of Second Language Writing, 40*, 1-11.

Lu, X. (2000). The influence of classical Chinese rhetoric on contemporary Chinese political communication and social relations. In D. R. Heisey (Ed.), *Chinese Perspective in Rhetoric and Communication* (pp. 3-23). Stamford, Connecticut: Ablex Publishing Corporation.

Lyons, J. (1995). *Semantics: An Introduction*. Cambridge: Cambridge University Press.

MacSwan, J., & Pray, L. (2005). Learning English bilingually: Age of onset of exposure and rate of acquisition among English language learners in a bilingual education program. *Bilingual Research Journal, 29*(3), 653-678.Maher, P. (2015). The role of 'that' in managing averrals and attributions in post-graduate academic legal texts. *English for Specific Purposes, 40*, 42-56.

DOI:10.6814/NCCU202100169

Malinowski, B. (1994 [1923]). The problem of meaning in primitive languages. In J. Maybin (Ed.), *Language and Literacy in Social Practice* (pp. 1-10). Clevedon: Multilingual Matters in association with The Open University.

Martin, J. & White, P. (2005). *The Language of Evaluation: Appraisal in English*. London: Palgrave.

Matthiessen, C. & Halliday, M. A. K. (1993). Systemic-Functional Grammar. In F. Peng & J. Ney (Eds.), *Current Approaches to Syntax* (pp. 2-30). London: Benjamins & Whurr.

Matalene, C. (1985). Contrastive rhetoric: An American writing teacher in China. *College English, 47*(*8*), 789-808.

Markkanen, R., & Schröder, H. (1989). Hedging as a translation problem in scientific texts. In C. Laurén & M. Nordman (Eds.), *Special Language: From Humans Thinking to Thinking Machines* (pp. 171-179). Clevedon: Multilingual Matters.

McCarthy, M. (2000). *Discourse Analysis for Language Teachers*. Cambridge: Cambridge University Press.

McEnery, T., & Hardie, A. (2011). *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

McGrath, L., & Kuteeva, M. (2012). Stance and engagement in pure mathematics research articles: Linking discourse features to disciplinary practices. *English for Specific Purposes, 31*(3), 161-173.

Miller, J. E., & Miller, J. (2011). *A Critical Introduction to Syntax*. London: Continuum.

Milton, J., & Hyland, K. (1999). Assertions in students' academic essays: A comparison of L1 and L2 writers. In R. Berry, B. Asker, K. Hyland, & M. Lam (Eds.), *Language Analysis, Description and Pedagogy* (pp. 147-161). Hong Kong: Hong Kong University of Science and Technology.

Mulholland, J. (1999). Email: Uses, issues and problems in an institutional setting. In F. Bargiela-Chiappini & C. Nickerson (Eds.), *Writing Business: Genres, Media and Discourses* (pp. 57-84). London: Longman.

DOI:10.6814/NCCU202100169

Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics, 10*(1), 1-35.

Myers, G. (1990). *Writing Biology: Texts in the Social Construction of Scientific Knowledge*. Madison, WI: University of Wisconsin Press.

Nakayama, S., & Dusenbury, J. (1984). *Academic and Scientific Traditions in China, Japan, and the West*. Tokyo: University of Tokyo Press.

Nash, W. (1990). The stuff these people write. In W. Nash (Ed.), *The Writing Scholar: Studies in Academic Discourse* (pp. 8-30). Newbury Park, CA: Sage.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*(2), 291-310.

Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes, 21*, 60-71.

Paraskevas, C. (2020). *Exploring Grammar Through Texts: Reading and Writing the Structure of English*. London: Routledge.

Parkinson, J. (2013). Adopting academic values: Student use of that-complement clauses in academic writing. *System, 41*(2), 428-442.

PCE. (2017). *PDF Converter Elite (Version 5.0.5)* [Computer Software]. Vancouver, BC: 0833266 B.C. Ltd.

Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist, 54*(9), 741-754.

Qiang, L. & Hua, W. (2010). Citation indexing formats of English academic literatures. *Journal of Modern Information*, *30*(9), 173-177.

Richards, K. & Skelton, J. (1991). How critical can you get? In P. Adams, B. Heaton & P. Howarth (Eds.), *Socio-Cultural Issues in English for Academic Purposes* (pp. 25-41). Basingstoke: Modern English Publications/British Council.

DOI:10.6814/NCCU202100169

Römer, U. (2009). Corpus research and practice: What help do teachers need and what can we offer? In K. Aijmer (Ed.), *Corpora and Language Teaching* (pp. 83-98). Amsterdam: John Benjamins.

Salager-Meyer, F. (1994). Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes*, *13*(2), 140-170.

Sanchez-Stockhammer, C. (2016). Punctuation as an indication of register: Comics and academic texts. In C. Schubert & C. Sanchez-Stockhammer (Eds.), *Variational Text Linguistics: Revisiting register in English* (pp. 139-167). Berlin: De Gruyter Mouton.

Schleppegrell, M. J. (2004). *The Language of Schooling: A Functional Perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schmid, H.-J. (2000). *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. Berlin: Walter de Gruyter.

Schmid, H. (2003). Probabilistic part-of-speech tagging using decision trees. In D. Jones, & H. Somers (Eds.), *New Methods in Language Processing* (pp. 154-164). London: Routledge.

Scollon, R. (2001) *Mediated Discourse Analysis: The Nexus of Practice*. London: Routledge.

Scollon, R. & Scollon, S. (1995). *Intercultural Communication.* Oxford: Blackwell.

Seely, J. (2007). *Oxford A–Z of Grammar and Punctuation.* Oxford: Oxford University Press.

Shi, L. (2003). Writing in two cultures: Chinese professors return from the West. *Canadian Modern Language Review*, *59*(3), 369-392.

Shore, S. (1993). A functional and social-semiotic perspective on language, context and text. In *SKY 1993*. Helsinki: Suomen kielitieteellinen yhdistys [The Finnish Linguistics Association].

Sigley, R. J. (1997). Text categories and where you can stick them: A crude formality index. *International Journal of Corpus Linguistics*, *2*(2), 199-237.

DOI:10.6814/NCCU202100169

Skelton, J. (1988). Comments in academic articles. In P. Grunwell (Ed.), *Applied Linguistics in Society* (pp. 98-108). London: GILT/BAAL.

Skelton, J. (1997). The representation of truth in academic medical writing. *Applied Linguistics, 18*(2), 121-140.

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, *328*(5977), 450-452.

Sowey, E., & Petocz, P. (2016). *A Panorama of Statistics: Perspectives, Puzzles and Paradoxes in Statistics*. Chichester: John Wiley & Sons.

Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing*, *32*: 17-35.

Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication*, *33*(2), 149-183.

Stubbs, M. (1996). *Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture*. Oxford: Blackwell.

Swales, J. M. (1987). Utilising the literatures in teaching the research paper. *TESOL Quarterly*, *21*(1), 41-68.

Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Swales, J. M. (1998). Textography: Toward a contextualization of written academic discourse. *Research on Language and Social Interaction*, *31*(1), 109-121.

Swales, J. M. (2004). *Research Genres: Explorations and Applications*. Cambridge: Cambridge University Press.

Swales, J. M., Ahmad, U. K., Chang, Y. Y., Chavez, D., Dressen, D. F., & Seymour, R. (1998). Consider this: The role of imperatives in scholarly writing. *Applied linguistics*, *19*(1), 97-121.

Swan, M. (2005). *Practical English Usage*. Oxford: Oxford University Press

DOI:10.6814/NCCU202100169

Tang, Y., Tseng, H., & Vann, C. (2020). Unwrap citation count, Altmetric Attention Score and Mendeley readership status of highly cited articles in the top-tier LIS journals. *Global Knowledge, Memory and Communication, Ahead-of-print*(Ahead-of-print).

Thetela, P. (1997). Evaluated entities and parameters of value in academic research articles. *English for Specific Purposes, 16*(2), 101-118.

Thompson, P. (2001). A pedagogically-motivated corpus-based examination of PhD theses: Macrostructure, citation practices and uses of modal verbs. Unpublished PhD thesis, University of Reading.

Thompson, P. (2012). Achieving a voice of authority in PhD theses. In K. Hyland & C. S. Guinda (Eds.), *Stance and Voice in Written Academic Genres* (pp. 119-133). London: Palgrave Macmillan.

Thompson, G., & Hunston, S. (2000). Evaluation: An introduction. In S. Hunston & G. Thompson (Eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse* (pp. 1-27). Oxford: Oxford University Press.

Thompson, G., & Ye, Y. (1991). Evaluation in the reporting verbs used in academic papers. *Applied Linguistics*, *12*(4), 365-382.

Thomson Reuters. (2019). *2020 Journal Citation Reports®*.

Tomlin, R. S. (2014). *Basic Word Order: Functional Principles*. Abingdon: Routledge.

Tran, L. T. (2006). Different shades of the collective way of thinking: Vietnamese and Chinese international students' reflection on academic writing. *Journal of Asia TEFL*, *3*(3), 121-141.

Tse, P. (2012). Stance in academic bios. In K. Hyland & C. S. Guinda (Eds.), *Stance and Voice in Written Academic Genres* (pp. 69-84). London: Palgrave Macmillan.

Tucker, P. (2003). Evaluation in the art-historical research article. *Journal of English for Academic Purposes*, *2*(4), 291-312.

Tweed, R. G., & Lehman, D. R. (2002). Learning considered within a cultural context: Confucian and Socratic approaches. *American Psychologist, 57*(2), 89-99.

van Maanen, J. (1995). *Representation in Ethnography*. Thousand Oaks, CA: Sage.

DOI:10.6814/NCCU202100169

van Noorden, R. (2015). Interdisciplinary research by the numbers. *Nature, 525*(7569), 306-307.

Vande Kopple, W. (2002). Metadiscourse, discourse, and issues in composition and rhetoric. In E. Barton & G. Stygall (Eds.), *Discourse Studies in Composition* (pp. 91-113). Cresskill, NJ: Hampton Press.

Vidal, M. C. Á. (1991). Towards a postmodern feminism? *Atlantis, 12*(2), 83-93.

Wang, H., Chen, B., & Chow, S. C. (2003). Sample size determination based on rank tests in clinical trials. *Journal of Biopharmaceutical Statistics, 13*(4), 735-751.

Wang, J. & Jiang, F. (2019). Research on stance expressions in academic writing by Chinese PhD students of science and engineering. *Foreign Language World, 3*, 23-31.

Weigle, S. C., & Friginal, E. (2015). Linguistic dimensions of impromptu test essays compared with successful student disciplinary writing: Effects of language background, topic, and L2 proficiency. *Journal of English for Academic Purposes, 18*, 25-39.

White, P. R. R. (2002). Appraisal – The language of evaluation and stance. In J. Verschueren, J.-O. Östman & J. Blommaert (Eds.), *Handbook of Pragmatics: 2002 Installment* (pp. 1-23). Amsterdam: John Benjamins.

White, P. (2003). Beyond modality and hedging: A dialogic view of the language of intersubjective stance. *TEXT - Special Issue on Appraisal, 23*(3), 259-284.

Winfield, B., Mizuno, T., & Beaudoin, C. (2000). Confucianism, collectivism and constitutions: Press systems in China and Japan. *Communication Law and Policy, 5*(3), 323-347.

Winter, E. O. (1977). A clause relational approach to English texts: A study of some predictive lexical items in written discourse. *Instructional Science,* 6, 1-92.

Winter, E. O. (1982). *Towards a Contextual Grammar of English: The Clause and its Place in the Definition of Sentence*. London: George Allen & Unwin.

DOI:10.6814/NCCU202100169

Wittenburg, K. (1987). Extraposition from NP as Anaphora. In J. Huck & A. Ojeda (Eds.), *Syntax and Semantics Vol. 20* (pp. 427–445). New York: Academic Press.

Wood, A. (2001). International scientific English: The language of research scientists around the world. In J. Flowerdew & M. Peacock (Eds.), *Research Perspectives on English for Academic Purposes* (pp. 71-83). Cambridge: Cambridge University Press.

Xie, C., & Teo, P. (2020). Tuning up the promotional volume: Comparing the About Us texts of top- and second-tier universities in China and America. *Journal of Pragmatics, 159*, 60-72.

Xu, H. L. (2011). Use of authorial stance markers in research discourse by Chinese advanced EFL learners: A corpus-based contrastive study. *Foreign Language Education, 32*(6), 44-48.

Yang, Y. (2013). Exploring linguistic and cultural variations in the use of hedges in English and Chinese scientific discourse. *Journal of Pragmatics*, *50*(1), 23-36.

Yin, Y. (2016). Western academic journals publishing mechanism revelation. *Editorial Friend, 10*, 30-37.

Zhang, F., & Zhan, J. (2020). Understanding voice in Chinese students' English writing. *Journal of English for Academic Purposes, 45,* Article 100844.

Zhang, J. (2011). Linguistic, ideological, and cultural issues in Chinese and English argumentative writings. *Journal of Language Teaching and Research*, *2*(1), 73-80.

Zhong, Y. (2006). *Research on Chinese Punctuation*. Beijing: China Federation of Literary and Art Circles Publishing House.

Zhang, Z. S. (2017). *Dimensions of Variation in Written Chinese*. London: Routledge.

Zuck, J. G. & Zuck, L. V. (1986). Hedging in news writing. In A.-M. Comu, J. Van Parjis, M. Delahaye, & L. Baten (Eds.), *Beads or Bracelets? How do We Approach LSP? Selected Papers from the Fifth European Symposium on LSP* (pp. 172-180). Oxford: Oxford University Press.

DOI:10.6814/NCCU202100169

APPENDIX

Journals represented in the corpus

Chemistry		Physics		Biology	
International	Local	International	Local	International	Local
1. *Nature Chemistry*	1. *Chinese Journal of Chemistry*	1. *Advances in Optics & Photonics*	1. *Light: Science & Applications*	1. *Trends in Plant Science*	1. *Molecular Plant*
2. *Accounts of Chemical Research*	2. *Science China Chemistry*	2. *Advances in Physics*	2. *Chinese Physics B*	2. *Nature Cell Biology*	2. *Journal of Integrative Plant Biology*
3. *Chem*	3. *Chinese Chemical Letters*	3. *Nature Photonics*	3. *Photonics Research*	3. *Molecular Biology and Evolution*	3. *Journal of Molecular Cell Biology*
4. *Journal of The American Chemical Society*	4. *Chemical Research in Chinese Universities*	4. *Reports on Progress in Physics*	4. *Chinese Journal of Physics*	4. *Nature Biotechnology*	4. *Frontiers in Biology*
5. *Chemical Science*	5. *Frontiers of Chemistry in China*	5. *Nature Physics*	5. *Frontiers of Physics*	5. *Trends in Cell Biology*	5. *Quantitative Biology*
Applied Linguistics		Law		Economics	
International	Local	International	Local	International	Local
1. *Modern Language Journal*	1. *Language and Semiotic Studies*	1. *Harvard Law Review*	1. *Frontiers of Law in China*	1. *Journal of Economic Perspectives*	1. *China & World Economy*
2. *Applied Linguistics*	2. *Concentric: Studies in Linguistics*	2. *Stanford Law Review*	2. *China Legal Science*	2. *Quarterly Journal of Economics*	2. *Frontiers of Economics in China*
3. *Journal of Linguistics*	3. *Taiwan Journal of Linguistics*	3. *Yale Law Journal*	3. *Tsinghua China Law Review*	3. *Journal of Finance*	3. *Frontiers of Business Research in China*
4. *Language Teaching Research*	4. *English Teaching & Learning*	4. *University of Pennsylvania Law Review*	4. *Peking University Transnational Law Review*	4. *Journal of Economic Growth*	4. *China Economist*
5. *TESOL Quarterly*	5. *Taiwan Journal of TESOL*	5. *Common Market Law Review*	5. *Asian Journal of WTO & International Health Law and Policy*	5. *Journal of Financial Economics*	5. *China Finance and Economic Review*

DOI:10.6814/NCCU202100169