

**Peoples' Friendship University of Russia Named After**

**Patrice Lumumba**

As a manuscript

**Han Hao**

**A comparative study of metadiscourse in Chinese EFL learners' and  
experts' academic writing**

5.9.8. Theoretical, applied, and comparative and contrastive linguistics

**THESIS**

for a PhD Degree

in Philology

**Scientific Supervisor:**

Candidate of philological sciences, Associate professor

Natalia M. Dugalich

**Moscow – 2025**

**Министерство науки и высшего образования РФ  
Федеральное государственное автономное  
образовательное учреждение высшего  
образования «Российский университет дружбы народов  
имени Патриса Лумумбы»**

На правах рукописи

**Хань Хао**

**Сравнительное исследование метадискурса в академическом письме  
китайских студентов, изучающих английский язык, и экспертов-  
лингвистов**

Специальность 5.9.8. Теоретическая, прикладная и сравнительно-  
сопоставительная лингвистика

**ДИССЕРТАЦИЯ**

на соискание учёной степени кандидата филологических наук

**Научный руководитель:**

кандидат филологических наук,

доцент Н.М. Дугалич

Москва – 2025

## Table of Context

<b>INTRODUCTION</b> .....	<b>5</b>
<b>CHAPTER I. METADISCOURSE</b> .....	<b>15</b>
1.1. Definitions of Metadiscourse .....	<b>15</b>
1.1.1. Metadiscourse in a broad sense .....	<b>16</b>
1.1.2. Metadiscourse in narrow sense.....	<b>20</b>
1.2. Different approaches to metadiscourse .....	<b>24</b>
1.2.1 J. Williams’ approach to metadiscourse .....	<b>24</b>
1.2.2. W. Vande Kopple’s approach to metadiscourse.....	<b>25</b>
1.2.3. A. Crismore’s approach to metadiscourse.....	<b>27</b>
1.2.4. K. Hyland’s approach to metadiscourse.....	<b>29</b>
1.2.5. A. Ädel’s approach to metadiscourse.....	<b>33</b>
1.2.6. N.K. Ryabtseva’s approach to metadiscourse .....	<b>36</b>
1.3. Metadiscourse research from different perspectives .....	<b>40</b>
1.3.1 Social interaction perspective on metadiscourse research.....	<b>40</b>
1.3.2 The functional perspective on metadiscourse research .....	<b>41</b>
1.3.3 Pragmatic perspective on metadiscourse research .....	<b>44</b>
1.3.4 A psycholinguistic perspective on metadiscourse research.....	<b>47</b>
1.4. Metadiscourse and rhetoric .....	<b>48</b>
1.4.1. The concept of rhetoric.....	<b>50</b>
1.4.2. Rhetorical hyping strategy.....	<b>51</b>
1.4.3. Hedging strategy.....	<b>53</b>
1.4.4. Authorial identity construction.....	<b>56</b>
<b>CHAPTER I CONCLUDING REMARKS</b> .....	<b>59</b>
<b>CHAPTER II. ENGLISH FOR ACADEMIC PURPOSES</b> .....	<b>63</b>
2.1. Definitions of EAP .....	<b>63</b>
2.2. Characteristics of English for academic purposes.....	<b>65</b>
2.3. Methodologies in English for academic purposes.....	<b>68</b>
2.3.1. Genre analysis .....	<b>68</b>

2.3.2. Academic discourse analysis.....	72
2.3.3. Corpus approach.....	74
2.3.4. Critical discourse analysis.....	76
2.3.5. Thematic analysis.....	78
<b>CHAPTER II CONCLUDING REMARKS.....</b>	<b>81</b>
<b>CHAPTER III. THE USE OF RHETORICAL STRATEGIES OF METADISOURSE IN THE ACADEMIC WRITING OF CHINESE EFL LEARNERS AND EXPERTS.....</b>	<b>83</b>
3.1. Research methodology.....	83
3.2 Rhetorical hype in academic writing: Investigation of certainty stance adverbs and maximizers.....	89
3.2.1 Investigation of certainty stance adverbs.....	90
3.2.2 Investigation of maximizer ‘fully’.....	102
3.3. Hedging strategy in academic writing: An investigation of lexical verbs.....	114
3.3.1. An investigation of speculative verbs.....	121
3.3.2. An investigation of quotative verbs.....	129
3.3.3. An investigation of sensorial verbs.....	137
3.4. Constructing authorial identity in academic writing: An investigation of self-mentions.....	143
<b>CHAPTER III CONCLUDING REMARKS.....</b>	<b>158</b>
<b>CONCLUSIONS.....</b>	<b>161</b>
<b>REFERENCES.....</b>	<b>167</b>

## INTRODUCTION

Metadiscourse, as a rhetorical framework, functions as a crucial tool for writers to provide cues for expressing their own stance within discourse, facilitating interaction and negotiation with readers regarding the viewpoints being presented [Hyland 2005]. This enhances reader engagement within the discourse, enabling a better understanding of the conveyed information and aiding in the acceptance of the writer's propositions, thereby achieving communicative goals.

The primary objective of English for academic purposes (EAP) writing is to showcase the writer's scholarly achievements, necessitating the persuasion of readers to believe in the writer's viewpoints. The appropriate use (individually or in combination) of metadiscourse resources (i.e. certainty stance adverb, booster, hedge and self-mention) contributes strategically (i.e. rhetorical hype strategy, hedging strategy and constructing authorial identity strategy) to this persuasion process, thus holding paramount importance in academic writing.

**The relevance of the study** can be elucidated through two main aspects. Firstly, it aligns with the broader general humanitarian interest in investigating discourse since the conditions, participants and rules of communication determine the choice of linguistic means and interact with the linguistic experience and personal characteristics of communicants. Secondly, it addresses the linguistic interest in the study of metadiscourse focusing on metadiscourse in the EAP writing of Chinese EFL learners and international journal experts.

**The degree of scientific development of the research problem.** This study undertook a comparative examination of EAP writing by Chinese EFL learners and international journal experts, focusing on the utilization of metadiscourse strategies.

The scientific bases for the present dissertation are the works dedicated to:

- *Genre analysis*: J. Swales (1990); T. Dudley-Evans (1994); K. Hyland (2005); L. Flowerdew (2004); M. Hoey (2001); E. Zanina (2017);
- *Critical discourse analysis*: N. Fairclough (1995); P. Baker et al. (2008); T.A. van Dijk (2008); V.I. Karasik (2020); E.N. Malyuga & B. Tomalin (2024); Y.A. Volkova & N.N. Panchenko (2024);
- *Academic discourse analysis*: A. Crismore (1993); U. Connor (1996); D. Biber et al. (1999); K. Hyland (2005); A. Ädel (2006); M.A.K. Halliday & C.M.I.M. Matthiessen (2014); S.W. Fitriati & N.M. Gayatri (2021); N.M. Dugalich & H. Hao (2024);
- *Corpus studies*: T. Johns (1986); A. Coxhead & P. Nation (2001); J. Sinclair (2004); D. Biber et al. (2006); K. Hyland (2008); T. Peredrienko & E. Balandina (2022);
- *Metadiscourse studies*: G. Bateson (1972); E. Goffman (1974); J. Rossiter (1974); A. Wierzbicka (1978); E. Keller (1979); D. Schiffrin (1980); J. Williams (1981); W. Vande Kopple (1985); N.K. Ryabtseva (1992); A. Crismore (1993); A. Mauranen (1993); K. Hyland (2005); A. Ädel (2006);
- *Metadiscourse and rhetoric*: K. Hyland & J. Milton (1997); R.A. Thabet (2018); & X. Ma (2019); A.Y. Almakrob (2020); F.K. Jiang (2023);
- *Studies on EAP writing*: R. Jordan (1997); K. Hyland (2002); R. Scarcella (2003); C.E. Snow & P. Uccelli (2009); N.X. Wei (2016); K. Hyland (2017); F. Jiang (2019); B.C. Lou (2022).

The **object** of the proposed dissertation research is metadiscourse in Chinese EFL learners' and experts' academic writing.

The **subject** of the present research is rhetorical hype, hedging strategy and authorial identity characteristics of metadiscourse in Chinese EFL learners' and experts' academic writing.

The **aim** of this study is to examine and delineate specific rhetorical strategies of metadiscourse in Chinese EFL learners' and experts' academic writing. To accomplish the goal, we pursued the following **objectives**:

1) to investigate the theoretical foundation of metadiscourse and EAP writing, and to define the boundaries of this thesis's exploration of metadiscourse;

2) to construct two corpora, namely, Chinese EFL learners' MA theses (referred to as CLMA\_C) and international linguistic journal articles (referred to as ILJA\_C), and prepare the methodological framework for subsequent comparative analysis;

3) to investigate and compare the employment of rhetorical hype strategies within the academic texts produced by Chinese EFL learners and experts, focusing on the usage of certainty stance adverbs and maximizers;

4) to explore and contrast the application of hedging strategies, with a particular focus on lexical verbs, in the academic texts produced by Chinese EFL learners and experts;

5) to examine and compare the construction of authorial identity within the academic texts produced by Chinese EFL learners and experts, with particular emphasis on the utilization of self-mentions;

6) to propose a metadiscourse analysis model, integrates corpus-genre methods in EAP research, and compare Chinese EFL/expert academic texts and offer learners structural (e.g. section templates), stylistic (e.g. conventional phrases), and pragmatic (e.g. hedging) tools, teachers corpus frameworks (e.g. concordance analysis) and developers contrastive resources (e.g. L1/L2 platforms, annotated corpora), advancing comparative linguistics and EAP pedagogy.

**The main hypothesis for the PhD defense** posits that through comparative analysis using corpora, Chinese EFL learners and experts exhibit metadiscourse features marked by identity attributes associated with rhetorical strategies in academic writing.

**Main provisions for the PhD defense:**

1. Academic writing embodies recognized genre characteristics, acknowledged by the academic community, while simultaneously being

culturally and identity-marked in the academic writing produced by Chinese EFL learners and experts.

2. In relation to rhetorical hype strategies in the academic texts produced by Chinese EFL learners and experts, differences emerge in the usage of certainty stance adverbs. Chinese EFL learners tend to favor active voice structures, whereas experts tend toward passive voice constructions.

3. Concerning hedging strategies, Chinese EFL learners tend to rely on impersonal constructions with hedging verbs more frequently than experts, aiming to uphold scientific objectivity.

4. In terms of authorial identity, experts demonstrate a preference for participant-oriented chunks when conveying evaluation and position, contributing to higher levels of academic persuasiveness and communicative effectiveness compared to Chinese EFL learners.

**The research data** employed in this thesis comprises 50 English linguistics MA theses (referred to as CLMA\_C) from 36 Chinese undergraduate universities, totaling 804,935 tokens, and 100 published articles from prominent international linguistic journals (referred to as ILJA\_C) in applied linguistics, totaling 802,490 tokens. These journals include *Applied Linguistics* (SNIP: 2.661), *English for Specific Purposes* (SNIP: 2.249), *International Journal of Corpus Linguistics* (SNIP: 1.211), *Journal of Pragmatics* (SNIP: 1.666), *Journal of Second Language Writing* (SNIP: 2.838), and *TESOL Quarterly* (SNIP: 2.359). Although the structural elements of articles in linguistic journals and MA theses are generally similar, the former tend to be longer. Thus, after consulting with academic experts in corpus linguistics and following the advice of my supervisor, it was decided to select 100 articles from influential international journals to ensure comparability, resulting in respective token counts of 804,935 and 802,490 for the two datasets. The research materials for this thesis include the entire papers, with the exception of the front page, Chinese abstract, list of tables and figures, table of contents, and appendix.



To conduct this study, we have integrated the following 5 **research methods**: corpus approach, comparative analysis, academic discourse analysis, critical discourse analysis and thematic analysis.

This research adopts **corpus approach** to examine metadiscourse features that facilitate rhetorical hype, hedging strategy, and authorial identity construction in 50 Chinese EFL learners' MA theses (CLMA\_C, 804,935 tokens) and 100 international journal articles (ILJA\_C, 802,490 tokens) from six applied linguistics journals (e.g., *Applied Linguistics*, *TESOL Quarterly*). Corpus design prioritizes comparability: CLMA\_C includes theses from 36 Chinese universities (2010–2013), while ILJA\_C comprises articles (2010–2013) with similar SNIP metrics (1.211–2.838) to ensure disciplinary rigor. Non-essential sections (e.g., front pages, appendices) were excluded to focus on core academic text (e.g., introductions, discussions).

**Comparative analysis** employs K. Hyland's (2005) interactional metadiscourse framework to further categorize linguistic features into certainty stance adverbs, maximizers, hedges and self-mentions. It is further complemented by quantitative methods, including normalized frequency analysis (per 1,000,000 words) and log-likelihood tests, to identify statistically significant differences between the examined groups.

**Academic discourse analysis** combines genre analysis (e.g., 'dissertation' VS 'research article') and qualitative examination of rhetorical strategies (e.g., how metadiscourse aid authorial identity construction).

To uncover the underlying ideological and power dynamics in metadiscourse, we integrated **critical discourse analysis (CDA)** into our study. CDA helped us interpret how language choices not only reflect but also reproduce broader cultural and institutional norms. For example, we examined the use of certainty stance adverbs (e.g., 'clearly') as a strategy to assert authority. In one instance, a journal article's frequent use of such certainty stance adverbs was interpreted as an attempt to convey expert certainty and align with dominant

academic standards. In contrast, fewer certainty stance adverbs in the Chinese theses suggested a different stance toward academic modesty and cultural norms regarding self-presentation.

Complementing our quantitative and qualitative linguistic analyses, **thematic analysis** was applied to the concordance lines generated from our corpora. This approach allowed us to identify recurring themes in the use of metadiscourse markers across different academic contexts. For example, a cluster analysis of engagement markers (such as “we believe,” “you can see”) revealed themes of collaborative knowledge construction. In both corpora, these markers were frequently associated with passages that aimed to involve the reader in the argument, though the exact lexical choices varied between the MA theses and the journal articles. This thematic insight deepened our understanding of how academic writers negotiate their authority and connection with the audience.

By integrating these five methods – each with concrete analytical examples – we obtain a multi-layered picture of metadiscourse practices. The corpus approach establishes a robust quantitative foundation; comparative analysis reveals statistically significant differences between corpora; academic discourse analysis and CDA provide interpretive depth regarding genre and ideology; and thematic analysis uncovers recurrent patterns across texts. Together, these methods ensure systematic, replicable insights into how metadiscourse reflects expertise levels and cultural-contextual norms in academic writing.

**The scientific novelty** of the proposed dissertation research resides in its pioneering comparative analysis of rhetorical hype strategy, hedging strategy, and the strategy of construction of authorial identity within the realm of metadiscourse in the academic texts produced by Chinese EFL learners and experts. Additionally, the dissertation describes the recognized genre characteristics inherent in academic writing, which are acknowledged by the academic community. Concurrently, it highlights how these characteristics are culturally and identity-marked in the academic writing of Chinese EFL learners

and experts. The dissertation represents the inaugural endeavor in research work focusing on metadiscourse resources and the analysis of their rhetorical strategies in a comparative context between Chinese EFL learners and experts.

### **Theoretical implications.**

This study contributes to academic discourse analysis by introducing three academic rhetorical strategies to scrutinise metadiscourse features embedded in scholarly texts, specifically focusing on rhetorical hype, hedging strategy, and authorial identity construction. Moreover, it broadens the analytical framework employed in EAP research by integrating corpus approach with genre analysis. Given that most corpus linguistic investigations of academic writing predominantly center on texts produced by native English-speaking scholars, this research extends the scope to encompass English academic compositions by Chinese MA linguistic students. Finally, the comparative dimension of the study is expanded to contrast the writing characteristics of Chinese English learners with those of proficient international academics.

### **Practical applications.**

This study equips EFL students, teachers, and materials developers with evidence-based tools grounded in corpus analysis and comparative insights. For students, it provides discipline-specific phrase lists (e.g., “This study demonstrates...”), alongside templates for critical sections (e.g., methods, discussions). Case studies contrast strong vs. weak texts to highlight pitfalls like L1 transfer errors or hedging misuse. Statistical data (e.g., metadiscourse frequency, collocation metrics) guide self-assessment.

Teachers gain frameworks for designing EAP courses, with activities like concordance analysis of hedging devices (“It is suggested that...”) and peer-review workshops using corpus benchmarks. Rubrics align with corpus-identified standards (e.g., “Effective introductions establish context in 2–3 sentences”). Discipline-specific modules address field conventions (e.g., self-mention markers in linguistic academic writing).

Materials developers leverage the corpus to create textbooks with real-world examples (e.g., annotated journal excerpts) and digital tools (e.g., interactive platforms comparing student academic writing to published texts). Specialized courses on contrastive rhetoric (e.g., English vs. Chinese conclusion structures) and corpus-assisted writing are supported by open-access data.

### **Approbation of the dissertation:**

The primary results and conclusions of this thesis were demonstrated amid *five* articles indexed in the Scopus international database, as well as in peer-reviewed journals listed by RUDN and VAK.

1. Dugalich, Natalia M. & Han Hao. (2024) Certainty Stance Adverbs in Chinese Linguistic Academic Writing: A Corpus-based Study. *RUDN Journal of Language Studies, Semiotics and Semantics*, 15(1), 248-261. <https://doi.org/10.22363/2313-2299-2024-15-1-248-261> (Scopus and VAK));

2. Dugalich N.M., Han H. (2024) Maximizers hyping in Chinese MA learners' and experts' academic discourse: An EUM-based study // *Litera*. № 3. P. 82-93. DOI: 10.25136/2409-8698.2024.3.70220 EDN: DYYVXQ URL: [https://en.nbpublish.com/library\\_read\\_article.php?id=70220](https://en.nbpublish.com/library_read_article.php?id=70220) (VAK);

3. Han H., Dugalich N.M. (2024) Self-mention in Chinese linguistic MA novices' and experts' academic writing: A corpus-driven investigation of 'we' // *Litera*. 2024. № 4. P. 182-194. DOI: 10.25136/2409-8698.2024.4.70516 EDN: TSJZST URL: [https://en.nbpublish.com/library\\_read\\_article.php?id=70516](https://en.nbpublish.com/library_read_article.php?id=70516) (VAK);

4. Han Hao. (2024) Lexical verbs hedging in Chinese linguistic academic writing // *Foreign languages in Tertiary Education*. 2024. №3. P. 94-98. DOI: 10.37724/RSU.2024.70.3.012. (VAK);

5. Han Hao. (2024). Proximity construction in Chinese linguistic MA novices' and experts' academic writing: Evidence from stance and engagement features. *Proceedings of «The 11th International Research Conference Topical*

*Issues of Linguistics and Teaching Methods in Business and Professional Communication» April. 333-336;*

6. Han Hao. (2024). Intercultural communication in Chinese linguistic MA novices' and experts' academic writing: An investigation of self-mention "we". *Proceedings of VII All-Russian Student Scientific and Practical Conference «Actual problems of intercultural communication» March. 389-392.*

Certain results from this research were also introduced at the subsequent **conferences:**

1. VII All-Russian Student Scientific and Practical Conference «Actual problems of intercultural communication». Moscow, RUDN, March 27th 2024;

2. The XI International Research Conference Topical Issues of Linguistics and Teaching Methods in Business and Professional Communication. Moscow, RUDN, April 17-28th 2024.

**Thesis structure.** For structural coherence, this paper is divided into distinct sections: an introduction, three chapters with respective concluding remarks, a final conclusion, and a comprehensive references section that includes 204 cited sources.

**Introduction** illuminates the relevance of the thesis, the degree of scientific development of the research problem. It delineates the object and subject of the research, alongside its purpose and objectives. Furthermore, the main hypothesis and provisions for the PhD defense are introduced, along with the research methodology encompassing the data and methods utilized in this thesis. Premises for theoretical and practical implementations of the research results are presented, emphasizing the scientific novelty of the thesis. Additionally, the structure of the proposed research paper is outlined, providing the full picture of the approbation of the results.

**Chapter I** examines the definitions, approaches, and perspectives related to metadiscourse. We can synthesize that, although the nomenclature of metadiscourse lacks uniformity and its boundaries are not sharply delineated,

overall, metadiscourse resources exhibit a continuum ranging from a broad to a narrow sense. The approaches to types is not absolute but rather contingent upon degrees of manifestation.

**Chapter II** reviews various definitions of EAP, highlighting its theoretical and pedagogical significance. It recounts the characteristics of EAP and provides detailed explanations of methodologies employed in EAP research, including genre analysis, discourse analysis, and corpus approaches.

**Chapter III**, focusing on the research results, analytic data, and discussion, begins by providing a detailed overview of the research methodology utilized in this thesis. This includes an examination of the research material, the three primary research methods utilized, and the main research framework. Subsequently, the chapter conducts separate investigations into rhetorical hype strategies, hedging strategies, and authorial identity construction in the academic texts produced by Chinese EFL learners and experts. This analysis is facilitated through an exploration of certainty stance adverbs, maximizers, lexical verbs and self-mentions.

## **CHAPTER I. METADISCOURSE**

Due to the significant implications of the metadiscourse concept for both oral and written communication, it has garnered attention from scholars in various disciplinary fields, encompassing diverse contexts. Simultaneously, owing to the inherent ‘fuzziness’ of this concept [Hyland 2005], linguists such as A. Ädel (2006), A. Crismore (1989), D. Schiffrin (1980), J. Williams (1981), K. Hyland (1998), and W. Vande Kopple (1985) have offered various definitions and classifications of metadiscourse. This chapter will provide an introduction to the concept of metadiscourse, covering its definitions, categorizations and research perspectives, thereby laying the groundwork for defining metadiscourse and developing the theoretical framework underpinning this investigation.

### **1.1. Definitions of Metadiscourse**

Researchers in the field of applied linguistics have extensively investigated the essential characteristics of metadiscourse and the operational mechanisms of metadiscourse markers within particular texts. This section aims to explore the intricacies involved in these investigations, thereby enhancing our understanding of metadiscourse.

Zellig Harris introduced the term ‘metadiscourse’ in 1959 to describe a method for understanding the practical use of language, providing an illustration that «metadiscourse refers to a writer’s or speaker’s attempts to guide a receiver’s perception and understanding of a text» [Ken Hyland 2005:3]. Scholars in different fields used different terms to express the concept of metadiscourse: ‘metacommunication’ by speech communication theorist J. Rossiter (1974); ‘frame’ by sociolinguists G. Bateson (1972), E. Goffman (1974); ‘meta-talk’ by sociolinguists D. Schiffrin (1980), L. Ragan and H. Robert (1981); ‘gambits’ psycholinguistic E. Keller’s (1979). These various terms reflect different

perspectives, understandings, and focal points in metadiscourse research, each making unique contributions to the examination of metadiscourse.

Currently, metadiscourse is often classified into broad and narrow categories based on its functions within discourse. Below, we will clarify the concept of metadiscourse from both broad and narrow perspectives.

### **1.1.1. Metadiscourse in a broad sense**

The broad tradition of metadiscourse is referred to be ‘integrative model’ [Hyland 2005]. Within this research tradition, textual interaction is considered the defining characteristic of metadiscourse, representing how authors or speakers signal their explicit presence within discourse. It encompasses linguistic resources used for discourse organization (i.e., textual functions), such as ‘in addition’, ‘finally’, ‘noted above’, ‘according to...’, ‘namely,’ as well as language resources conveying the author’s attitude, evaluation, and stance towards the proposition content and the reader’s attitude (i.e., interpersonal functions), emphasizing the interactive relationship between the writer and the reader [Hyland 2005], such as ‘might,’ ‘definitely,’ ‘surprisingly,’ ‘we,’ ‘consider,’ among others. Both types of resources aid readers or listeners in organizing, interpreting, and evaluating discourse information. Scholars employing this approach include: A. Crismore (1989), J. Williams (1981), K. Hyland (1998), W. Vande Kopple (1985), among others. Next, we will introduce these scholars’ definitions of metadiscourse.

J. Williams applied metadiscourse to the field of writing, defining it as «writing about writing, whatever does not refer to the subject matter being addressed» [Williams 1981:211–212]. His perspective posited metadiscourse as a stylistic feature adept at illustrating discourse organization, enhancing structural discernibility, sourcing opinions and facts, discerning the author’s intent, and assisting readers in constructing discourse understanding. J. Williams also



emphasized the utilization of metadiscourse across various genres, noting its infrequent use in instructional manuals, technical manuals, scientific writing, and legal texts, but frequent use in literary criticism, narratives, and personal letters. He observed its most common usage in argumentative writing, as authors frequently direct statements towards viewpoints and readers' understanding in discourse.

Following J. Williams' definition of metadiscourse, W. Vande Kopple divides discourse into two tiers: «On one level we supply information about the subject of our text; on this level we expand propositional content. On the other level, the level of metadiscourse, we do not add propositional material but help our readers organize, classify, interpret, evaluate, and react to such material» [Vande Kopple 1985:83]. Moreover, metadiscourse was conceptualized as «discourse that people use not to expand referential material but to help their readers connect, organize, interpret, evaluate, and develop attitudes towards that material» [Vande Kopple 1997:2]. To put it differently, those linguistic elements that do not contribute propositional content but signal the author's presence.

A. Crismore (1989) contends that all linguistic communication comprises two tiers of discourse: the 'primary discourse' level, which encompasses 'referential' and 'informational' meanings facilitating the expression of ideas and experiences, and the 'secondary discourse' or 'metadiscourse' level, which includes propositional meanings, textual meanings and interpersonal meanings [Crismore 1989]. In this regard, the metadiscourse level serves the primary discourse level, making the author's intent more easily understood and accepted by readers. Based on this, A. Crismore developed a rhetorical model of metadiscourse studies and regard its function as «a social, rhetorical instrument which can be used pragmatically to get things done» [Crismore 1989:4]. Functioning as a reader-guiding linguistic mechanism, metadiscourse is employed by authors to direct rather than merely inform readers, involving in subtle Socratic dialogues about motives, aims, connotations, and stances

[Crismore and Farnsworth 1989]. A. Crismore conducted an analysis on the cultural and gender variations in the application of metadiscourse within student compositions in both the USA and Finland, emphasizing that metadiscourse pertains to «the linguistic material in texts, whether spoken or written, that does not add anything to the propositional content but that is intended to help the listener or reader organize, interpret, and evaluate the information given» [Crismore 1993:40]. Its effectiveness is based in «how readers should read, react to, and evaluate what they have written about the subject matter. Writers convey their personality, credibility, considerateness of the reader, and relationship to the subject matter and to readers by using certain devices in their texts» [Crismore 1993:39-40]. Metadiscourse guides readers in the discourse, helping them better understand the discourse and the author's viewpoints.

The aforementioned function-based definitions of metadiscourse distinctly separate metadiscourse from propositional content, emphasizing that metadiscourse is independent of propositional content. Some scholars have raised objections to this. R. Mao, through examples, demonstrates that metadiscourse in certain contexts also has truth values and belongs to propositions [Mao 1993]. Therefore, he advocates for blurring the boundaries between basic discourse and metadiscourse. R. Luukka also raises doubts, stating that if metadiscourse is defined as elements that do not add to the propositional meaning of discourse, then removing metadiscourse without altering the genuine meaning of discourse would be unreasonable and defines metadiscourse as «non-propositional elements of texts, but, instead, take a more functional approach» [Luukka 1994:78]. Furthermore, R. Luukka proposes the concept of 'contextual metadiscourse' in his research on metadiscourse in the oral and written versions of five Finnish-language papers submitted to academic conferences [Luukka 1994]. He argues that authors and readers evaluate the communicative scene and the discourse as a scene, emphasizing more the interactional function of metadiscourse, while giving less attention to the textual function.

K. Hyland is among the scholars who adopt a functional classification, employing the concept metadiscourse «to discuss those aspects of the text which explicitly refer to the organization of the discourse or the writer's stance towards either its content or the reader» [Hyland 2004:109]. It functions as a broad term encompassing diverse cohesive and interpersonal elements, linking and structuring content in alignment with both the author's intent and the conventions of the discourse community [Hyland 1998]. Through skillful use of metadiscourse, authors can elevate an opaque text into one that is coherent and accessible to readers, situating it within specific contexts to convey the author's persona, authority, reader responsiveness, and the author's connection with the content [Hyland 2000]. Consequently, K. Hyland categorizes metadiscourse as a functional component. Furthermore, he critiques the drawbacks of segregating metadiscourse from propositional content, contending that a text's meaning hinges on the amalgamation of all its elements, blurring the distinction between propositional and non-propositional aspects. K. Hyland delineates three fundamental principles of metadiscourse: (1) its distinction from the propositional facets of discourse; (2) its embodiment of writer-reader interactions within the text; and (3) its restriction to relationships intrinsic to the discourse [Hyland 2005]. In response to deficiencies in prior metadiscourse definitions, K. Hyland introduces an interpersonal model, asserting that «Metadiscourse is the cover term for the self-reflective expressions used to negotiate interactional meanings in a text, assisting the writer (or speaker) to express a viewpoint and engage with readers as members of a particular community» [Hyland 2005:38].

According to the extensive definitions of metadiscourse offered by the scholars mentioned earlier, it becomes apparent that metadiscourse pertains to elements within discourse explicitly tied to discourse structure or the author's perspective, demonstrating the author's attentiveness to the reader. The focus of a comprehensive approach lies in the overt representation of the author within the discourse, encompassing the expression of attitude or assessment towards the

discourse, and the illustration of discourse structure. Metadiscourse serves readers by aiding in the organization, interpretation, and assessment of information within the discourse. Thus, metadiscourse fulfills a persuasive rhetorical role by aiming to shape the reader's comprehension of the discourse.

### **1.1.2. Metadiscourse in narrow sense**

A narrow approach to examining metadiscourse is often labeled as 'non-integrative' or the 'reflexive model' [A. Ädel 2010; A. Wierzbicka 1978; Mauranen 1993]. This research approach excludes interpersonal factors and confines the domain of metadiscourse to linguistic components used solely to achieve discourse functions, emphasizing its role in discourse organization.

A. Wierzbicka contends that discourse consists of statements about things and statements about discourse itself, wherein speakers utilize metadiscourse to elucidate certain aspects of the discourse itself. Statements about things are bound together with the discourse about the discourse using 'lines' [Wierzbicka 1978]. In a sense, these 'lines' can stitch together statements about things into a tightly cohesive whole, endowing them with strong coherence. These metadiscourses imbue discourse with relevance, yet they themselves are heterogeneous entities. When separating these heterogeneous components, we can use 'scissors' to cut out the part belonging to the 'meta-voice' in the double-voiced discourse, namely, the metadiscourse.

D. Schiffrin is a sociologist frequently cited in metadiscourse research, focusing on metadiscourse in spoken discourse and using the term 'meta-talk' to discuss talk about talk in conversation [Schiffrin 1980]. In her study of conversations, she found that speakers use meta-talk resources such as 'Well', 'That's what I meant', 'I am arguing that' and 'I'm telling you that' to organize and evaluate conversations. This self-directed evaluation differs from the organization of information by speakers and allows them to present themselves

as ‘animators’, engaging actively in the generation of discourse. Additionally, D. Schiffrin categorizes meta-talk into two main types: ‘organizational brackets’ and ‘evaluative brackets’ [Schiffrin 1980]. ‘Bracket’ here refers to metalinguistic clauses that mark the boundaries between discourse units (conversation blocks with inherent structures). Organizational brackets can initiate or terminate conversation blocks (e.g., ‘There are five points here’), while evaluative ones identify the speaker’s stance on the spoken discourse (e.g., ‘That’s interesting’). D. Schiffrin also notes that «meta-talk functions on a referential, informational plane when it serves as an organizational bracket, and on an expressive, symbolic plane when it serves as an evaluative bracket» [Schiffrin 1980:231]. She also acknowledges the multifunctionality of some meta-talk resources, such as ‘That’s my opinion’, which serves as organizational bracket when signaling the end of the speaker’s viewpoint and evaluative bracket when expressing the speaker’s attitude.

A. Mauranen (1993) is a prominent figure in this research tradition and, like D. Schiffrin, she also employs the term ‘metatext’ to describe metadiscourse phenomena [Mauranen 1993]. She argues that the functions of metatext extend beyond propositional content and are primarily used for text-organizing. She emphasizes the explicit evaluation of the discourse by the author, exemplified by phrases such as ‘Let us consider the consequences of the theory outlined earlier’, ‘We shall divide such factors into three categories as follows...’, ‘The book is...’, ‘in the upcoming segment...’, ‘in addition...’ and so on. She also highlights that the defining characteristic of metatext is its reference to the discourse or the writing process itself, rather than external elements. However, A. Mauranen’s understanding of metatext overlooks two crucial factors related to writing: the writer and the reader [Mauranen 1993].

A. Ädel (2006) also understands metadiscourse in terms of ‘self-reflexivity’, arguing that when we communicate, we do not only talk about the world and ourselves, but also use language to talk about talking. Humans can

explicitly present themselves as ‘experiencers’ and ‘communicators’, enabling them to evaluate both the topics being discussed and the communicative context. Put simply, speakers of all human languages can reflexively use language to assess language-related matters [Ädel 2006].

A. Ädel defines metadiscourse as «text about the evolving text, or the writer’s explicit commentary on her own ongoing discourse» [Ädel 2006:20]. This concept includes two components: ‘metatext’ and ‘writer-reader interaction’. ‘Metatext’ denotes reflexive linguistic expressions that refer to the discourse or its linguistic structures, while ‘writer-reader interaction’ refers to the envisioned presence of the author or reader within the discourse. The linguistic items to which metadiscourse refers reveal the presence of the author and reader in the discourse, enabling discourse organization and evaluation. Metadiscourse takes various forms, ranging from «morphemes, single word forms, phrases, clauses, to strings of sentences» [Ädel 2006: 2].

A. Ädel’s (2006) contributions to metadiscourse research primarily encompass the following elements: (1) introducing the ‘reflexive triangle’ to delineate the interactions among ‘author’, ‘text’, and ‘reader’, emphasizing the roles of discourse participants; (2) outlining three facets of reflexivity: how authors self-refer, how authors engage with readers, and how authors discuss their own discourse. The initial aspect involves explicit mentions of oneself by authors, including first-person pronouns (both singular and plural) and impersonal constructions. The second aspect relates to relational markers, such as imperatives and inclusive ‘we’, through which authors explicitly establish relationships with readers. The third aspect is realized through metatext or discourse clues left by authors to guide readers in research papers; (3) summarizing four characteristics of metadiscourse: it is a fuzzy category with boundaries that are difficult to delineate; it is a functional category, implemented at the lexical-grammatical level through various forms and structures; expressions of metadiscourse may be multifunctional, serving several linguistic

functions simultaneously; it depends on context (i.e. ‘context-dependent’), meaning that determining whether a lexical-grammatical means implements metadiscourse function requires close examination of context.

Through an examination of previous research, it is found that scholars who adopt a narrow perspective on metadiscourse often use the term ‘metatext’ rather than ‘metadiscourse’. This tendency can be partially explained by the distinction between ‘text’ and ‘discourse’. The term ‘text’ narrowly denotes the written documentation of a communicative event or action, whereas ‘discourse’ is generally perceived as a broader term encompassing the entirety of the communicative event or action. [Brown and Yule 1983: 5–12]. Therefore, ‘metadiscourse’ is generally used as a broad term encompassing interpersonal and textual factors, while ‘metatext’ is used as a narrow term focusing solely on textual factors.

This dissertation approaches the concept of metadiscourse from broad perspective, utilizing an operational definition grounded in K. Hyland’s interpersonal framework. There are several justifications for this choice. Firstly, K. Hyland’s definition was developed within the context of academic discourse, which aligns with the present research’s focus on academic writing—a key component of that discourse domain [Hyland 2005]. Secondly, his classification is structured around how linguistic elements convey an author’s epistemic stance toward a given proposition [Hyland 1998]. This perspective resonates with the scope of this study, which examines metadiscourse usage among Chinese EFL writers and expert authors. Lastly, K. Hyland’s methodological approach to metadiscourse in academic writing frequently incorporates corpus-based techniques, mirroring the analytical methods employed in this research [Hyland 2021]. Consequently, this study adopts K. Hyland’s interactional metadiscourse framework for both operational definition and classification.

## **1.2. Different approaches to metadiscourse**

Understanding metadiscourse methods involves not only examining its definition but also its classification. While the notion of metadiscourse is easy to grasp, «establishing its boundaries is difficult» [Li 2018: 14]. Throughout the past three decades, various classification models of metadiscourse have been proposed by A. Ädel (2006); A. Crismore (1993); K. Hyland (2005); W. Vande Kopple (1985). The differences in metadiscourse classification models may be attributed to the following factors: (1) the ambiguity of metadiscourse definitions and the diversity of functions due to differences in disciplines and research perspectives, as elaborated upon in the preceding section; (2) the development of metadiscourse models frequently relies on the categorization and analysis of extensive corpora, and variations in metadiscourse classification models may be attributed to the specific corpora utilized by researchers [Mu 2010]. For example, A. Crismore's metadiscourse classification model is based on an examination of sociological textbooks, while K. Hyland & P. Tse used graduate theses as their corpus [Crismore 1985; Hyland & Tse 2004]. Overall, these classifications tend to become more comprehensive, clear, and simplified. Within this section, our focus will be to discuss several influential metadiscourse classification models.

### **1.2.1 J. Williams' approach to metadiscourse**

J. William is a trailblazer in metadiscourse research, exerting significant influence on subsequent scholars in the field. In his seminal work [William 1981], he posits that metadiscourse serves as a means for authors to directly or indirectly instruct readers on how to interpret their ideas. The examples of metadiscourse he provides encompass methods of altering subjects and expressing certainty in viewpoints. He regards metadiscourse as stylistic variants, categorizing them into three main types, each further subdivided into paired subclasses.



(1) *Hedges and emphatics*. The former indicate the author's uncertainty regarding assertions, leaving room for discussion and speculation (e.g., 'perhaps', 'maybe', 'tend to', 'assume'), while the latter are used to emphasize the author's assertions, demonstrating the credibility of the claims (e.g., 'clearly', 'evidently', 'show', 'obviously').

(2) *Sequencers and topicalizers* are both employed to guide readers within the discourse. Sequencers, serving as transitional words, aid authors in organizing the discourse, ensuring its coherence and cohesion (e.g., 'for one thing', 'to start with', 'next', 'finally'). Topicalizers, on the other hand, announce transitions in topics (e.g., 'to come back to', 'with regard to', 'in terms of', 'to move on').

(3) *Attributors and narrators* serve distinct functions within discourse. Attributors are utilized to denote the source of opinions and facts to readers (e.g., 'according to', 'claim', 'believe', 'demonstrate'). On the other hand, narrators provide a narrative of the author's thought process, often using first-person pronouns as subjects (e.g., 'What I am trying to say', 'I think', 'I believe').

J. Williams' model distinguishes metadiscourse into paired stylistic devices – hedges with emphatics, sequencers with topicalizers, and attributors with narrators – to guide readers' interpretations. This approach foregrounds the author's role in signaling uncertainty or certainty and in organizing discourse. However, its dichotomous pairings tend to oversimplify language use by not fully accommodating the multifunctionality and context-dependent nature of many expressions. As academic texts become increasingly complex, such rigid classifications might struggle to capture nuanced shifts in tone and structure, thereby limiting its applicability across varied discursive contexts.

### **1.2.2. W. Vande Kopple's approach to metadiscourse**

Initially, W. Vande Kopple comprehensively categorized metadiscourse from a functional viewpoint, building on J. Williams' investigations to classify it into seven distinct types. [Vande Kopple 1985; Williams 1981].

(1) *Text connectives*: These metadiscourse resources facilitate readers' smooth comprehension of discourse and its appropriate construction. They are further categorized into six subtypes: 'words and phrases that indicate sequences' (e.g., 'initially', 'subsequently', 'thirdly'); 'words and phrases denoting logical and time-related connection' (e.g., 'however', 'nevertheless', 'as a consequence', 'at the same time'); 'reminders about material presented earlier in texts' (e.g., 'as was mentioned in Section Two'); 'announcements of material appearing later in text' (e.g., 'as will be demonstrated in the subsequent section'); 'declarations regarding the material that is about to be presented' (e.g., 'My current aim is to elaborate on the concept that') and 'topicalizers' (e.g., 'for example', 'there are', 'with respect to', 'concerning') [Vande Kopple 1985:83–85].

(2) *Code glosses*: These metadiscourse resources assist readers in understanding the precise meaning of specific parts of discourse through methods such as definitions, annotations, supplementary explanations, and specifying scope, as seen in phrases like 'in other words' and 'that is'.

(3) *Illocution markers*: These words explicitly identify the author's speech acts in specific parts of discourse. Authors use phrases and clauses like 'I propose that', 'in conclusion', 'we assert that', 'I promise to' and 'for example' to implement acts such as hypothesizing, summarizing, declaring, promising, and giving examples [Vande Kopple 1985:83–85].

(4) *Validity markers*: These words indicate the author's judgment of the possibility or truth value of proposition content, expressing the author's viewpoint on the credibility of proposition material. Validity markers consist of three types: hedges (e.g., 'perhaps', 'may', 'might', 'seem', 'to a certain extent'), emphatics (e.g., 'clearly', 'undoubtedly', 'it is obvious that'), and attributors (e.g., 'according to Einstein').

(5) *Narrators*: These metadiscourse resources primarily serve to notify the audience of the sources of information, as seen in phrases like 'according to James', 'Mrs. Wilson announced that', and 'the principal reported that'.

(6) *Attitude markers*: These words reveal the author's attitude towards proposition content, as seen in expressions like 'unexpectedly', 'it is intriguing that' and 'it is concerning to observe that'.

(7) *Commentary*: Commentaries actively involve readers in an implicit exchange with the author, exemplified by phrases such as 'most of you will oppose the idea that', 'you might wish to read the last chapter first', 'you will probably find the following material difficult at first' and 'my friends.'

W. Vande Kopple expands metadiscourse into seven functional types – including text connectives, code glosses, illocution markers, validity markers, narrators, attitude markers, and commentary—that address both textual and interpersonal dimensions. This granular framework offers detailed insight into the mechanisms that foster coherence and engage readers. Yet, its complexity is a double-edged sword: the overlapping functions of certain markers can lead to ambiguous categorization, making consistent application across diverse academic texts challenging. This intricacy, while comprehensive, may render the model less accessible for streamlined analytical purposes.

### **1.2.3. A. Crismore's approach to metadiscourse**

Cross-cultural research on the utilization of metadiscourse in argumentative texts by American and Finnish university students was conducted by A. Crismore, exploring how gender factors influence the usage of metadiscourse [Crismore 1993]. Building upon W. Vande Kopple's classification system, A. Crismore retained the classification of metadiscourse into textual metadiscourse and interpersonal metadiscourse, while reorganizing and rearranging their subcategories based on specific pragmatic functions [Crismore 1993].

In the domain of textual metadiscourse, A. Crismore adapted W. Vande Kopple's model by omitting 'narrators' and 'temporal connectives,' but keeping

‘sequencers,’ ‘reminders,’ ‘logical connectives,’ and ‘topicalizers.’ They also rebranded the primary category from ‘text connectives’ to ‘textual markers’ for better accuracy in scope. Additionally, they introduced a new category named ‘interpretive markers,’ which encompassed ‘code glosses,’ ‘announcements,’ and ‘illocution markers.’ Within this category, ‘validity markers’ were further divided into three subcategories: ‘hedges,’ ‘certainty markers,’ and ‘attributors,’ while ‘attitude markers’ and ‘commentary’ continued as subcategories. Moreover, due to the dual function of ‘narrators’ and ‘attributors’ in informing or persuading readers about information sources, they were merged into one category. Therefore, A. Crismore revised model of metadiscourse is as follows:

(1) *Textual Metadiscourse*: This category of metadiscourse elements is employed to delineate the structural framework of discourse within the text. It comprises two subcategories:

Textual markers, including logical connectives (e.g., ‘and’, ‘but’, ‘therefore’, ‘in addition’, ‘however’, ‘because’, ‘which’), sequencers (e.g., ‘first’, ‘second’), reminders (e.g., ‘as we noted earlier’), and topicalizers (e.g., ‘well’, ‘now the question arises’, ‘in regard to’).

Interpretive markers, including code glosses (e.g., ‘namely’, ‘for example’, ‘what I mean is’), illocution markers (e.g., ‘I state again that’, ‘I plead with you’), and announcements (e.g., ‘in the next section’).

(2) *Interpersonal metadiscourse*: This variety of metadiscourse tools is used to facilitate interaction between readers, texts, and authors, comprising five subcategories:

Hedges (e.g., ‘can’, ‘could’, ‘may’, ‘might’, ‘would’, ‘I think’, ‘I suppose’ etc.)

Certainty markers (e.g., ‘certainly’, ‘I know’, ‘really’ etc.)

Attributors (e.g., ‘Einstein claimed that’, ‘I find it surprising that,’ etc.)

Attitude markers (e.g., ‘I hope’, ‘unfortunately’, ‘most importantly’, ‘doubtfully,’ etc.)

Commentary (e.g., ‘you may not agree that’, ‘as a colleague’, ‘think about it’, ‘let us’ etc.).

A. Crismore (1993) further developed the metadiscourse classification model of W. Vande Kopple (1985) by not only revising the categories and subcategories but also including punctuation marks. Specifically, in certain contexts, if the content following punctuation marks such as colons, commas, underscores, and parentheses is explanatory, it can be categorized as code glosses within interpretive markers, exemplified by phrases like ‘That is the commitment Wilkins is making: indulging.’

Building on W. Vande Kopple’s model, Crismore reconfigures metadiscourse into two overarching categories – textual and interpersonal – while merging similar functions and introducing interpretive markers. This reorganization simplifies analysis and aligns more closely with pragmatic communication in academic writing. Nonetheless, the consolidation may sacrifice some of the nuanced distinctions present in prior frameworks, potentially overlooking subtle differences between functions like narrating and attributing. In contexts where fine-grained analysis is critical, such reduction may diminish the model’s explanatory power, suggesting that further refinement could be beneficial.

#### **1.2.4. K. Hyland’s approach to metadiscourse**

K. Hyland, based on A. Crismore’s metadiscourse classification model, analyzed 28 academic texts from four disciplines: microbiology, economics, astrophysics, and applied linguistics, from a pragmatic perspective [Hyland 1998; Crismore 1993]. He categorized metadiscourse into two main types: textual metadiscourse and interpersonal metadiscourse, and further subdivided them into more specific subcategories. K. Hyland’s metadiscourse classification model is outlined below:

(1) *Textual Metadiscourse*: This category of language resources constructs persuasive and coherent discourse by establishing explicit interpretations of propositions and relationships between propositions and other discourses, enabling readers to understand the author's intentions. It includes five subcategories:

Logical connectives: These primarily involve conjunctions, indicating progression, causation, contrast, etc., aiding readers in understanding the pragmatic relationships between viewpoints, such as 'in addition', 'but', 'therefore', 'thus', 'and' etc.

Frame markers: They provide explanatory framework information for longer discourse segments, explicitly marking discourse boundaries or indicating structural elements, such as sequence ('initially', 'subsequently'), discourse stages ('to summarize', 'in conclusion'), discourse objectives ('I contend here', 'my intention is'), topic shifts ('By the way', 'Anyhow').

Endophoric markers: They refer to expressions within the discourse that indicate references to other parts of the text ('see Table 2', 'as noted above'), making additional conceptual material explicit and helping readers grasp the author's argumentative intent.

Evidentials: Similar to endophoric markers, evidentials denote the origin of information derived from other discourses, establishing intertextuality, showcasing related research within the field, and guiding readers' interpretations of the discourse, such as 'according to X', '(Y, 1990)', 'Z states' etc.

Code glosses: Authors provide additional information through explanations, comparisons, expansions, etc., reflecting their expectations of readers' knowledge and comprehension abilities, such as 'for example', 'put differently' etc.

(2) *Interpersonal metadiscourse*: This category of metadiscourse resources signals the author's attitude towards proposition information and the reader, constructing author-reader relationships and subjective negotiation of

propositions. Essentially, interpersonal metadiscourse is interactive and evaluative, expressing the author's persona in communicative acts. It comprises five subcategories:

Hedges: They express the author's uncertainty in presenting and evaluating proposition information, such as 'possible', 'might', 'perhaps', etc.

Emphatics: Used to emphasize the author's certainty and the strength of propositions, such as 'it is obvious', 'definitely', 'of course' etc. The balance between hedges and emphatics reflects the author's assertion of propositions, catering to reader's face needs and rhetorical community norms.

Attitude markers: Authorial stance towards discourse message can be elucidated through phrases like 'interestingly', 'regrettably', 'I concur' and so forth.

Relational markers: They are explicit means that writers directly engage readers, aiming to capture their attention or involve them as active participants in the discourse by using phrases like 'honestly', 'it's important to note' or 'one can see'.

Person markers: These markers indicate the author's level of presence in discourse which is assessed by how often first-person pronouns or possessive adjectives are used, such as 'I', 'we', 'my', 'mine', 'our' etc. [Hyland 1998].

While presenting this metadiscourse classification model, K. Hyland acknowledged its limitations: in actual language use, some linguistic items can serve multiple pragmatic functions simultaneously. The demarcation line within metadiscourse and proposition content is not always clear-cut, and the relationship between metadiscourse categories implementing textual function and interpersonal function is often not binary. For instance, contrastive conjunctions like 'but' and 'however,' while introducing positive responses to assertions, can also convey politeness and serve interpersonal function.

Due to the constraints encountered in the pragmatic application of K. Hyland's (1998) taxonomy of metadiscourse, he later refined and enhanced this

framework, ultimately formulating an interpersonal model of metadiscourse, which will be demonstrated explicitly in section 3.1. K. Hyland's (2005) interpersonal model reconceptualizes metadiscourse by dividing it into two broad categories – interactive and interactional markers – that work in tandem to shape academic writing. In the interactive domain, markers focus on the text's architecture and clarity and help to guide the reader through the text. For example, transition markers (e.g., 'however', 'therefore') signal logical relations between propositions, while frame markers (e.g., 'firstly', 'in conclusion') delineate the organization and progression of arguments. Endophoric markers (e.g., 'as noted above', 'see Figure 1') guide readers to related text segments, and evidentials (e.g., 'according to Smith (2003)') attribute information to external sources, reinforcing credibility. Additionally, code glosses (e.g., 'in other words', 'that is') rephrase or clarify ideas to enhance reader comprehension. In contrast, the interactional category emphasizes the writer's engagement with the audience, involving the reader in the text. Hedges (e.g., 'perhaps', 'might') introduce caution or nuance, while boosters (or emphatics, e.g., 'clearly', 'indeed') assert confidence in claims. Attitude markers (e.g., 'surprisingly', 'importantly') reveal evaluative stances, self-mentions (e.g., 'I', 'we') personalize the text and establish authorial presence, and engagement markers (e.g., 'you may notice', 'note that') directly invite reader participation. Together, these subcategories not only structure the discourse but also create a dynamic writer–reader relationship, ensuring that arguments remain both coherent and persuasively engaging.

This dissertation employs this categorization because it organizes linguistic features based on their interactive and interactional role in expressing an author's engagement with and response to a particular assertion [Hyland 1998]. This theoretical framework corresponds with the objectives of this study, which investigates how Chinese EFL writers and professional scholars utilize interactive and interactional metadiscourse. Furthermore, since this taxonomy was originally formulated within the sphere of academic communication [Hyland



2005], it is particularly relevant to the present research, which concentrates on academic writing as a fundamental aspect of that communicative field.

### **1.2.5. A. Ädel's approach to metadiscourse**

Drawing from R. Jakobson's language functions, A. Ädel advanced a 'reflexive model' within the categorization of metadiscourse, categorizing it into 'personal metadiscourse' and 'impersonal metadiscourse'. [Ädel 2006; Jakobson 1985]. Personal metadiscourse directly addresses the writer or reader in discourse using personal pronouns (e.g., 'I', 'we', 'you') or nouns (e.g., 'the writer', 'dear reader'). Within the same discourse community, this explicit referencing considers individuals as active participants, indicating that the addressed entities are the writers and potential readers of the discourse. In contrast, impersonal metadiscourse implicitly refer to discourse participants, and its realization may involve passive voice or other non-personal structures. Impersonal metadiscourse is categorized into four types: (1) phoric markers, such as 'below', 'following', 'before', 'above', 'back to', 'begin', 'last', (2) Citations to the text, such as 'essay', 'paragraph', 'text', 'phrase', 'sentence', 'word', (3) code glosses, such as 'briefly', 'i.e.', 'mean', 'namely', and (4) discourse labels, such as 'mention', 'put', 'say', 'question', 'such as', 'conclude', 'aim', 'stress'.

A. Ädel also presents four criteria for identifying metadiscourse within the reflexive model:

(1) 'Explicitness': Requires the explicit statement of what is referred to in the discourse world.

(2) 'World of discourse': Focus must be on the ongoing discourse, rather than worldly behaviors or phenomena outside of other discourses.

(3) 'Current discourse': Metadiscourse elements should pertain to the present discourse instead of external ones, thereby illustrating intertextuality.

(4) «‘Writer qua writer’ and ‘reader qua reader’: For a linguistic unit to be considered metadiscourse, it must be used to indicate the roles of both the speaker/writer and audience/reader in the world of discourse, contrasting with real individuals in the real world» [Ädel 2006: 27].

This reflexive model proposed by A. Ädel offers several advantages:

Firstly, this model exhibits a high level of theoretical precision by presenting clear criteria, which limit the breadth of the ambiguous notion of metadiscourse and differentiate it from other phenomena such as stance and participation. Secondly, by emphasizing the reflexivity of metadiscourse (in contrast to non-propositional content). Thirdly, it circumvents the issue of propositional element within metadiscourse. Lastly, by encompassing both authors and readers, it situates the concept of metadiscourse within specific contexts.

A. Ädel combined two research methods, qualitative and corpus-based comparative analysis, to investigate the use of personal metadiscourse in 30 university lectures and 130 graduate research papers. The aim was to explore the academic discourse functions of spoken and written metadiscourse, which led to the promotion of a new categorization system for metadiscourse. [Ädel 2010]. This categorization system, as presented by A. Ädel, comprises four categories encompassing 23 discourse functions:

(1) *Metalinguistic Comments*

Repairing: For instance, ‘I’m sorry,’ ‘I didn’t intend to speak that aloud.,’ etc.

Reformulating: Such as ‘if you’ll permit me,’ ‘adjust it slightly,’ ‘or...if you want,’ etc.

Commenting on linguistic form/meaning: For example, ‘I’m not sure how to phrase this exactly, but...,’ ‘if you will...,’ etc.

Clarifying: For instance, ‘I am not suggesting that...,’ ‘To clarify, I should mention that...,’ etc.

Managing terminology: Such as ‘What we imply by this is that...,’ ‘I define... as...,’ etc.

## (2) Discourse Organization

Introducing topic: For example, ‘in this paper I explore...,’ etc.

Delimiting topic: Such as ‘I will focus on...,’ etc.

Adding to topic: For instance, ‘we might add that...,’ etc.

Concluding Topic: For example, ‘we’ve now covered...,’ ‘we conclude that...,’ etc.

Marking aside: Such as ‘I’d like to take a brief detour here...,’ etc.

Enumerating: For instance, ‘We will discuss...first,’ ‘I have two objections against...,’ etc.

Endophoric marking: Such as ‘based on these reference points,’ ‘it is evident that...,’ etc.

Previewing: For example, ‘as I discuss below...,’ ‘as will become apparent subsequently,’ etc.

Reviewing: Such as ‘as I discussed above...,’ etc.

Contextualizing: For instance, ‘I have addressed only briefly...,’ etc.

## (3) Speech act labels

Arguing: For example, ‘I am postulating that...,’ ‘I argue that...,’ etc.

Exemplifying: Such as ‘I intend to explore the polite case as a case study...,’ etc.

Other examples of speech act labeling include phrases such as and ‘I am proposing that...’ etc.

## (4) Addressing the audience

Facilitating understanding/communication: Such as ‘you know what I mean?’ ‘I didn’t catch that...,’ etc.

Managing audience discipline: For example, ‘May I have your attention, please?’ ‘Could we please have a moment of quiet?’ etc.

Anticipating the audience's response: Such as 'you may be thinking that...,' 'you may be curious about...,' etc.

Managing the message: For instance, 'I trust you found these materials insightful,' 'I have endeavored to provide the reader with...' etc.

Imaging scenarios: Such as 'suppose I say that...,' 'imagine the following situation...,' etc.

A. Ädel (2010) presents this classification model as an extension and refinement of A. Ädel (2006), allowing for a comprehensive description of both oral and written metadiscourse. It is crucial to emphasize that this model is proposed based solely on academic discourse, and its applicability to informal conversations requires further validation. Additionally, this study only focuses on personal metadiscourse, leaving room for further research on impersonal metadiscourse categories.

A. Ädel's reflexive model, inspired by Jakobson's language functions, separates metadiscourse into personal and impersonal dimensions, with clearly defined criteria emphasizing explicitness and contextual relevance. By further detailing functions into metalinguistic comments, discourse organization, speech act labels, and audience addressing, the model achieves theoretical precision. Nonetheless, its stringent criteria and focus on academic discourse may constrain its adaptability to informal or cross-genre contexts. The heavy emphasis on explicit reflexivity, while insightful for academic texts, might overlook subtler, implicit metadiscursive signals present in more diverse communicative settings.

#### **1.2.6. N.K. Ryabtseva's approach to metadiscourse**

Since N.K. Ryabtseva proposed the term *метатекст* [metatekst] 'metatext', metadiscourse has emerged as a central focus of research within the realm of Russian discourse analysis. The linguistic expressions in metadiscourse are varied, utilizing a range of distinctive discourse markers within the discourse,

often characterized by fixed or semi-fixed structures. Overall, the main sources of metadiscourse in academic discourse include modal particles (e.g., *даже* [dazhe] ‘even’, *ведь* [ved’] ‘after all’), insertions (e.g., *наконец* [nakonets] ‘finally’, *во-первых* [vopevyh] ‘firstly’), conjunctions and connectors (e.g., *однако* [odnako] ‘however’, *следовательно* [sledovatel’no] ‘hence’), adverbs (e.g., *поэтому* [poretomu] ‘therefore’, *напротив* [naprotiv] ‘opposite’), modal verbs (e.g., *возможно* [vozmozhno] ‘may’), word combinations (phrases or structural combinations) (e.g., *по этой причине* [po etoy prichine] ‘for this reason’, *иными словами* [inymi slovami] ‘in other words’), and other commonly used discourse markers. These metadiscursive elements in discourse serve to link propositions or different discourses, thereby creating coherent discourse and indicating semantic relationships between sentences and overarching units. While in certain instances, the logical connections within the information are evident, metadiscourse serves the purpose to make the connections between sentences more concrete [Shvedova 1980].

Academic discourse not only entails the objective presentation of research findings by authors but also involves a subjective reconstruction of the research activity. Consequently, some metadiscursive elements in discourse express the author’s viewpoints, attitudes, and subjective evaluations of propositions. In addition to common discourse markers, a considerable portion of metadiscourse in academic discourse consists of specific types, including:

(1) Mental performatives (*ментальные перформативы* [mental’nye performativy]): The concept of mental performatives was first introduced by N.K. Ryabtseva. She stated in her study of scientific discourse that speech acts can accomplish not only social (normative) actions but also mental actions [Ryabtseva 1992]. Differentiating from the performative utterances discussed by J.L. Austin, which aim to change social space, N.K. Ryabtseva referred to statements accomplishing mental actions as mental performatives (e.g., *Отвлечемся от параметра P* [otvlechemsya ot parametra] ‘Let’s abstract from

parameter P’). N.K. Ryabtseva suggested that a mental performative, as a type of statement, meant that the speaker, by uttering it, completed a mental operation in the process of inference [Ryabtseva 1992]. Mental performatives serve as explicit markers of cognitive operations, indicating and determining the corresponding steps of inference, thereby constructing the content of discourse.

In sentences containing mental activity verbs, these verbs are characterized by: employing the first person (often in the plural first person); using future or present tense forms (*ниже представим* [nizhe predstavim] ‘let’s present below’, *особо подчеркну* [osobo podcherknu] ‘I will particularly emphasize’, *выделяем* [vydelyaem] ‘we highlight’); co-occurring with modal or evaluative words (*надо отметить* [nado otmetit’] ‘it should be noted’, *интересно отметить* [interesno otmetit’] ‘it is worth noting’); co-occurring with adverbial verb phrases (*соглашаясь с этим мнением, укажем* [soglashayas’ s etim mneniem, ukazhem] ‘in agreement with this opinion, let’s indicate’; *резюмируя проведенное выше обсуждение, отметим* [resumiruya provedennoye vyshe obsuzhdenie, otmetim] ‘summarizing the above discussion, let’s note’); employing impersonal forms (*предполагается, что...* [predpolagaetsya, chto...] ‘it is assumed that...’; *утверждается, что...* [utverzhdetsya, chto] ‘it is asserted that...’); adopting hypothetical forms (*можно было бы утверждать* [mozhno bylo bi utverzhdat’] ‘one could assert’). Although different grammatical forms are employed for these activity verbs, it is important that all these forms of activity have not changed in meaning, and the moment of speaking coincides with the moment when the action represented by the verb is completed [Ryabtseva 1992].

(2) Descriptive (indirect) mental performatives: According to N.K. Ryabtseva, mental performatives can be used for description, where mental performative verbs are used not to perform actions but to describe. These sentences lack the external formal characteristics of performative sentences and are either used for substitution (e.g., replace the normative performative sentence

*ниже мы приводим примеры в фонематической записи* [nizhe my privodim primery v fonematicallyeskoj zapisi] ‘Below we provide examples in phonemic transcription’ with *ниже примеры приводятся в фонематической записи* [nizhe primery privodiatsya v fonematicallyeskoj zapisi] ‘Below examples are given in phonemic transcription’) or to establish mutual connections between sentences in discourse (e.g. *как уже отмечалось выше, что* [kak uzhe otmechalos’ vyshe, chto] ‘as previously noted that’; *выше мы уже указывали, что* [vyshe my uzhe urfzyvali, chto] ‘We have already indicated above that’).

(3) Fixed nominal groups consisting of general scientific nouns (e.g., *вопрос* [vopros] ‘question’, *гипотеза* [gipoteza] ‘hypothesis’, *синтез* [sintez] ‘synthesis’, *анализ* [analiz] ‘analysis’, *система* [sistema] ‘system’, *функция* [funktsiya] ‘function’, *понятие* [ponyatie] ‘concept’) and mental performative verbs (e.g., *подвергать рассмотрению* [podvergat’ rassmotreniyu] ‘examine’, *проводить сравнение* [provodit’ sravnenie] ‘compare’). These general scientific nouns typically denote abstract mechanisms of scientific cognition activities, often common terms whose metadiscursive function is only clarified in specific usage contexts.

(4) Typical formulas composed of general scientific vocabulary, such as *по моему мнению* [po moemu mneniyu] ‘in my opinion’, *не вызывает сомнения* [ne vyzyvayet somneniya] ‘there is no doubt’, *наша гипотеза состоит в том* [nasha gipoteza состоit v tom] ‘our hypothesis is that’, *в заключительной главе делается вывод* [v zaklyuchitel’noy glave delaetsya vyvod] ‘the final chapter concludes’, *как следует из приводимых ниже примеров* [rfr sleduet iz privodimyh nizhe primerov] ‘as follows from the examples below’, *статья посвящена какой проблеме* [statya posviashchena kakoy probleme] ‘the article is devoted to what problem’, *речь идет о чем* [rech idet o chem] ‘what are we talking about’, etc. The construction of scientific information heavily relies on these commonly employed standard formulas in academic discourse. They

encompass words or phrases of different semantic and grammatical categories, independent of specific object domains, characterized by fixity.

N.K. Ryabtseva's approach, rooted in Russian discourse analysis, foregrounds metadiscourse through fixed or semi-fixed expressions such as modal particles, insertions, connectors, and mental performatives. This framework effectively captures culturally and linguistically specific markers that contribute to text cohesion and authorial subjectivity. However, its reliance on formulaic expressions can be limiting; the model may not fully account for the fluid and evolving nature of metadiscourse in dynamic academic writing. As language usage varies significantly across disciplines and contexts, the fixed nature of this categorization might impede its broader applicability beyond the studied genre.

### **1.3. Metadiscourse research from different perspectives**

Through the review of the previous sections, we can observe that scholars have proposed various definitions and classification models of metadiscourse. The following part emphasizes the distinct perspectives of scholars from fields such as communication theory, sociolinguistics, psycholinguistics, pragmatics, cognitive linguistics, and rhetoric on the study of metadiscourse.

#### **1.3.1 Social interaction perspective on metadiscourse research**

Studies on metadiscourse from a social interaction standpoint mainly investigate its role in human communication and the social factors that affect how verbal communication is produced and interpreted. Early scholars who conducted metadiscourse research from this perspective include J. Rossiter (1974), a discourse interaction theorist, social linguists such as E. Goffman (1974) and D. Schiffrin (1980), among others.



J. Rossiter (1974) used the term ‘metacommunication’ to refer broadly to metadiscourse. He posited that all discourse about discourse constitutes additional comments on interaction, with metadiscourse aiding in the understanding of the intended message by explicitly identifying the author’s intentions within the discourse. J. Rossiter distinguished two types of metadiscourse: metadiscourse about one’s own discourse and metadiscourse about others’ discourse (evaluating other researchers’ literature). He also highlighted that metadiscourse can assist participants in exploring, interpreting, and clarifying the interaction process, suggesting that it is a process factor that can influence readers’ emotional responses to verbal communication. This perspective highlights the importance of metadiscourse in the construction of discourse.

Social linguist E. Goffman (1974) regarded metadiscourse in human interaction as ‘frames’, providing readers or listeners with framing information to interpret human interaction discourse. Seen from this angle, metadiscourse adheres to specific scene and genre norms, requiring mental effort from both parties in the linguistic community to apply and understand metadiscourse. D. Schiffrin (1980), a sociolinguist, saw metadiscourse as ‘metatalk’ in verbal interaction, helping listeners grasp the speakers’ motives. According to her argument, metadiscourse allows speakers to govern fundamental discourse as a means for speakers to influence how listeners were comprehended the whole discourse. Additionally, social linguists J. Ragan and R. Hopper (1981), like D. Schiffrin (1980), referred to the explicit forms of metadiscourse in human interaction as ‘metatalk’.

### **1.3.2 The functional perspective on metadiscourse research**

Research on metadiscourse from the functional perspective benefits from the insights of M.A.K. Halliday, the founder of systemic functional linguistics,

regarding the functional aspects of language, a viewpoint widely adopted by metadiscourse researchers W. Vande Kopple (1985), K. Hyland (1998). M.A.K. Halliday perceives language as a symbolic framework that emphasizes the application of language within social contexts. Within the framework of systemic functional linguistics, language is described as having the action potential and meaning potential, namely, 'can do', 'can mean' and 'can say' [Hyland 1973:43]. According to M.A.K. Halliday (1985), meaning is a form of behavior that can be manifested through optional lexical and grammatical potential. The three metafunctions of language, which guide language behavior choices, include the ideational function, the interpersonal function, and the textual function. They are detailed below: «The ideational function pertains to the use of language to represent experience and ideas. This concerns perceptions of the world and our own consciousness. The interpersonal function refers to the use of language to encode interaction, allowing us to engage with others, to take on roles and to express and understand evaluations and feelings. The textual function points to the use of language to organize the text itself, coherently relating what is said to the world and to readers» [Hyland 2005:26]. M.A.K. Halliday (1973) states that these function act as 'mediator', allowing language users to express their personality and emotions regarding the conceptual content of discourse and evaluate the propositional content. Additionally, it allows language users to intervene in various forms of social interaction with other participants in communicative contexts, providing them with ways to guide readers in interpreting the propositional content. Choices within the textual system are related to the organizational function of language in discourse, enabling speakers to organize discourse content in a coherent manner. The process of making choices is dynamic, and its ultimate purpose is to construct meaning at different levels. The overall system of language meaning is constituted by these three metafunctions, which are the primary means for participants in discourse to achieve their communicative purposes.

M.A.K. Halliday's concept of metafunctions laid the theoretical foundation for metadiscourse research, and many metadiscourse researchers, including A. Crismore and R. Farnsworth (1990), K. Hyland (1998), and W. Vande Kopple (1985), utilize metafunctions in their studies. Viewed functionally, propositional meaning serves the ideational role, whereas metadiscourse directly addresses the interpersonal and textual roles of language [Vande Kopple 1985]. Nonetheless, according to M.A.K. Halliday, every clause enacts these three functions simultaneously, and metadiscourse is not exempt from this rule. As noted by A. Crismore, «grammatical and lexical factors can simultaneously convey ideational meaning and interpersonal meaning, or simultaneously convey textual meaning and interpersonal meaning» [Crismore 1989:65] Therefore, in addition to serving interpersonal and textual meaning, metadiscourse also plays a role in constructing propositional content.

A. Crismore put forward an extensive rhetorical model for researching metadiscourse, which aligns with M.A.K. Halliday's functional view of language, thereby also regarded as a functional approach to metadiscourse. A. Crismore (1989) suggests that any discourse, whether oral or written, is inherently rhetorical, involving three elements: author, reader, and subject, with the author communicating with the reader on a specific topic. Both metadiscourse and basic discourse can serve three rhetorical purposes: informational, persuasive, and expressive, closely related to M.A.K. Halliday's propositions about interpersonal and textual functions of language.

Drawing from R. Jakobson's functional theory of language, A. Ädel employs a functional approach in examining metadiscourse, highlighting the involvement of discourse participants and using the concept of the 'reflexive triangle' to explore the connections among text, author, and reader. [Ädel 2006; Jakobson 1960]. Other researchers adopting a functional perspective on metadiscourse include L. Dillon (1981) and R. Luukka (1992). L. Dillon proposes that the development of discourse behavior is intricately linked to the utilization

of metadiscourse, which divides specific writing-reading scenes into two levels, one focusing on the author providing readers with information about the topic and expanding the indicative content, and the other focusing on the author's writing itself and the readers' reading of the content, which is related to metadiscourse. R. Luukka does not define metadiscourse as «non-propositional factors of discourse» but instead 'adopts more of a functional approach' to study metadiscourse» [Luukka 1992:78].

### **1.3.3 Pragmatic perspective on metadiscourse research**

Metadiscourse possesses 'inherent pragmatics' [Cao and Hu 2014:18]. As pointed out by A. Fuertes-Olivera, metadiscourse can assist users in «achieving their specific purposes in accordance with two basic language principles: cooperation and least effort» [Fuertes-Olivera 2001:1292]. According to K. Hyland, metadiscourse plays a crucial role as a pragmatic construct, illuminating how authors influence readers' understanding of discourse and their own stance as authors. [Hyland 1998; Li 2018]. The pragmatic perspectives in metadiscourse research include Speech Act Theory [Beauvais 1989], the Principle of Cooperation [Kumpf 2000], and Relevance Theory [Bu 2014; Ifantidou 2005].

J. Beauvais (1989) incorporated metadiscourse research into the framework of Speech Act Theory, developing a speech act model of metadiscourse. Within the context of various speech act studies [Austin 1962; Searle 1969], she defined metadiscourse as «illocutionary force indicators that identify expositive illocutionary acts» [Beauvais 1989:15]. In this context, the employment of metadiscourse markers can underscore the discourse acts that the author intends to perform at particular stages of the discourse. J. Beauvais categorized metadiscourse into primary and secondary categories, each with canonical and partially explicit forms [Li 2018]. This model uses first-person pronouns to describe explicitly off-locutionary acts (e.g., 'I should point out') and

second or third person pronouns to describe implicitly illocutionary acts (e.g., ‘Brown states that’, ‘You may consider’). Besides, J. Beauvais’s speech act model can be employed to discourse analysis, as illustrated by the example ‘I believe that tax reform is necessary.’ This example contains a performative sentence using the present tense of the verb ‘believe’ in the first person, conveying illocutionary force through the clause ‘I believe that’. However, not all metadiscourse can be reflected through explicit performative sentences; For example, consider the following sentence: ‘It has been observed that large-scale testing is increasing.’ There is no explicit performer of the illocutionary act. A. Crismore (1989) pointed out limitations in this model, suggesting it does not apply to broader language functions and rhetorical principles. Although J. Beauvais’s speech act model of metadiscourse has limitations, metadiscourse appears to be a type of illocutionary act within the context of Speech Act Theory.

Prior to J. Beauvais, other scholars had also noted the relevance of metadiscourse to speech act, with W. Vande Kopple (1980) noting that metadiscourse is often used to mark progression of basic discourse and demonstrate the author’s position on the basic discourse and then shape the reader’s attitude towards the basic discourse, which concerns of speech act. According to A. Crismore (1985), metadiscourse serves to highlight the performative nature of speech acts, identifies the reader as an active participant and marks the presence of the author. A. Crismore (1989) integrates the three metafunctions of language with speech act theory to elucidate the roles of metadiscourse and the actions it accomplishes. Both scholars are aware of the multiple communicative functions performed by metadiscourse and the author’s goals in communication, but they fail to fully examine the possibilities offered by speech acts.

Research on interpreting metadiscourse based on P. Grice’s Cooperative Principle is relatively limited [Grice 1975]. P. Kumpf (2000) extended metadiscourse into the visual domain, integrating the consistency of visual

metadiscourse with P. Grice's maxim of relation, suggesting that readers can anticipate the coherence of items within discourse [Kumpf 2000]. He emphasized that consistency and coherence describe the visual coherence of discourse, meeting the modernist demands for order and unity. Moreover, P. Kumpf suggested that extending the Cooperative Principle to the domain of writing is feasible, as it can enhance the functionality of metadiscourse, allowing authors to adopt the role of 'cooperative communicators'. R. Abdi conducted a qualitative examination of metadiscourse strategy selection, in conjunction with K. Hyland's metadiscourse model and the concept of P. Grice's Cooperative Principle in research papers, constructing a Cooperative Principle model that operates within the usage of metadiscourse strategies [Abdi 2010; Hyland 2005]. They argued that communication within discourse communities through academic papers can be seen as cooperative behavior shaped by implicit and explicit conventions.

E. Ifantidou challenges scholars such as J. Williams, W. Vande Kopple and A. Crismore, who interpret metadiscourse as non-propositional, rhetorical, stylistic, and secondary aspects of language [Ifantidou 2005; Vande Kopple 1985]. Drawing on authentic linguistic data from linguistic and medical textbooks, E. Ifantidou offers a reinterpretation of the attributes of metadiscourse through the lenses of both semantic and pragmatic viewpoints, employing D. Sperber and D. Wilson's relevance theory framework. [Ifantidou 2005; Sperber and Wilson 1995]. At the semantic level, metadiscourse facilitates the sharing of propositional content within discourse; at the pragmatic level, it is an indispensable element for effective interpretation of academic discourse. J. Bu examines how metadiscourse manifests in academic lectures through the lens of relevance-adaptation theory. This perspective suggests that analyzing the construction of relevance-adaptation patterns can illuminate the pragmatic description and functions of metadiscourse in academic settings [Bu 2014]. M. Aguilar (2008) explores how metadiscourse is employed in engineering academic discourse through the lens of D. Sperber and D. Wilson's relevance theory,

exploring the motivations behind the use of metadiscourse. When interpreting discourse using relevance theory, both discourse factors and non-linguistic factors such as the speaker's intention, social context, and cognitive context are involved.

Although K. Hyland does not give an explicit discursive model of metadiscourse, he consistently treats metadiscourse as a discursive construct [Hyland 1998]. In discussing the phenomenon of metadiscourse in academic discourse from a pragmatic perspective, he points out that academic discourse should combine illocutionary acts with perlocutionary effects, and that metadiscourse identifies the communicative intentions of the author in presenting the meaning of a proposition. It provides a means of discursive presupposition, helping the reader to process the discourse, encode connections between ideas, and order the material in a way that is appropriate to persuade the potential author.

#### **1.3.4 A psycholinguistic perspective on metadiscourse research**

Psycholinguists also investigate metadiscourse in human communication, focusing on the strategies speakers employ to organize their thoughts and establish communication steps. In the field of psycholinguistics, speakers utilize various strategies to organize their speech wordings and adjust their discourse acts, intending to capture audience's attention at key moments, aid in interpreting intentions, and facilitate social interaction. These strategies are presented in fixed or semi-fixed forms, referred to as 'formulaic sequences' by A. Wray (2000). This perspective explores the psychological activities and usage strategies of speakers or writers, providing a new dimension for understanding human communication.

Representative of this research perspective is E. Keller (1979), who refers to metadiscourse as 'gambits', used to introduce topic transitions within conversations or to prepare listeners for the arrival of the next turn in

conversation. ‘Gambits’ are signal or information words used by speakers as conversation strategies to accomplish specific communicative tasks. E. Keller (1979) suggests that speakers use ‘gambits’ to perform the following functions: presenting topics, structuring conversation turns, indicating their awareness of information, viewpoints, knowledge, emotions, or planned actions, and confirming whether communication has been conveyed to listeners. He also summarizes four main characteristics of ‘gambits’: semantic framing, social signaling, state-of-consciousness signaling, and communicative signaling. These constitute the basis of semantic, social, psychological, and communicative control signals used by speakers. The ‘gambits’ used by speakers in conversation construct conversation strategies to achieve specific communicative goals. The form of ‘gambits’ used also requires cognitive effort from the speakers.

From a structural perspective, some metadiscourse markers (such as ‘strictly speaking,’ ‘if you will,’ ‘of course,’ ‘in addition to’) exhibit fixed or semi-fixed forms. Psycholinguists use terms such as ‘prefabricated phrases’, ‘formulaic sequences’, ‘conventionalized language’, ‘collocations’ and ‘chunks’ to name them. Due to the short-term nature of memory, the use of these information units in chunk form can enhance the efficiency of information reception by listeners [De Carrico and Nattinger 1998:92]. The analysis of phrase sequences in this manner can further elucidate the characteristics of metadiscourse.

#### **1.4. Metadiscourse and rhetoric**

Metadiscourse is intrinsically linked to the communicative intentions of both authors and speakers, enabling them to embed their perspectives, assessments, and interpretative judgments within a text. This mechanism facilitates the structuring and refinement of concepts while taking into account the anticipated reception by the audience. As a strategic linguistic tool,



metadiscourse serves to engage readers, establish textual coherence, convey degrees of epistemic certainty, and steer comprehension, ultimately aligning with persuasive functions. Historically rooted in rhetorical traditions dating back to ancient Greece, metadiscourse underpins three fundamental persuasive appeals:

Logos is reinforced through explicit connections between concepts and arguments, enhancing coherence.

Ethos, or credibility, is established as metadiscourse reflects the author's expertise and authoritative stance.

Pathos, or emotional engagement, is invoked when the writer acknowledges the audience's perspective and tailors discourse to resonate with their expectations.

This section explores the interplay between metadiscourse and rhetorical techniques as mechanisms of persuasion and further clarify the discussed rhetorical strategies in this dissertation, that is, rhetorical hyping strategy, hedging strategy and authorial identity construction.

In this dissertation, metadiscourse functions as a broad rhetorical framework, encompassing various strategies that shape academic writing [Hyland 2005]. The three rhetorical strategies examined in this study can be understood as concrete manifestations of metadiscursive strategies frequently employed in scholarly texts. In other words, they represent specific applications of the 4 metadiscourse resources, either individually or in combination. For instance, the rhetorical hype strategy involves the use of certainty stance adverbs and maximizers, both subcategories of metadiscourse, to amplify claims within academic writing [Jiang 2023]. Similarly, the hedging strategy is realized through the deployment of hedges, another subcategory of metadiscourse, which serve to introduce caution and flexibility in scholarly arguments [Vass 2017]. Lastly, in the constructing authorial identity strategy, self-mention markers, one additional subcategory of metadiscourse, are commonly utilized to establish the writer's

presence and facilitate interaction with the audience [Jiang and Ma 2019; Jiang and Wang 2021].

#### **1.4.1. The concept of rhetoric**

The concept of rhetoric has evolved significantly over time, once holding a prominent place as a key academic discipline (Ong, 1983). Fundamentally, rhetoric refers to the craft of persuasion, dealing with arguments that lack definitive proof. In earlier periods, the term often carried negative associations, implying deceptive manipulation. However, in contemporary discourse analysis and studies of written communication, rhetoric has regained scholarly importance. As A. Mauranen (1993b) notes, rhetoric has been rediscovered not only as a tool for refining communication but also as a method for examining discourse across various disciplines.

Classical rhetorical principles remain integral to modern studies of spoken communication and writing theories, many of which still reflect Aristotle's foundational work on persuasion (Erickson, 1974; Furley & Nehamas, 1994). Aristotle's rhetoric is a cornerstone text that systematically investigates the mechanics of persuasion, distinguishing rhetoric from dialectic by focusing on practical argumentation. He posited that persuasion is achieved when an audience perceives a claim as credible, necessitating strategies that effectively communicate truth. Aristotle identified three essential elements in persuasion: the speaker, the audience, and the argument itself. Additionally, he emphasized three critical factors in constructing an argument: the persuasive techniques employed, linguistic choices, and structural organization.

These principles remain central in modern writing pedagogy, particularly in guiding students on argumentation, language selection, and discourse structuring. Of particular relevance to metadiscourse research are Aristotle's three modes of persuasion:

Ethos: the credibility and authority of the speaker or writer. While initial credibility may stem from expertise, reputation, or status, it must also be reinforced throughout the discourse. Contemporary perspectives view ethos as an emergent property of the interaction between the writer and the reader, shaped by textual engagement [Hauser 1986].

Pathos: the appeal to emotions, which considers audience characteristics such as background knowledge, interests, and socio-demographic factors.

Logos: the logical structure of the argument, encompassing its organization, complexity, and use of supporting evidence.

Although these modes of persuasion function independently, they typically work in unison to enhance persuasive impact. In relation to metadiscourse, these modes of persuasion are evident in different textual strategies: logos is reflected in explicit logical connections between ideas in arguments, ethos is established through demonstrations of writer's expertise and credibility, and pathos is conveyed by aligning discourse with the readers' expectations and values. The following discussion will explore how metadiscourse manifests three modes of persuasion within rhetorical hyping strategy, hedging strategy and authorial identity construction within academic writing.

#### **1.4.2. Rhetorical hyping strategy**

Academic papers serve as a fundamental medium for disseminating and exchanging scientific knowledge, with their credibility rooted in the logical rigor of research findings and their disciplinary significance. Traditionally, scholarly writing has been perceived as an impartial representation of experimental data and scientific analysis, thereby excluding rhetorical devices, which are often seen as contrary to the essence of scientific discourse [Liu et al. 2021]. However, empirical investigations within the sociology of scientific knowledge, which closely examine scientists and their research processes, reveal that scientific facts

are not merely discovered but are instead constructed through rhetorical framing to establish them as universally valid truths [Gilbert & Mulkay 1984]. As a result, academic texts inherently reflect the author's stance and broader social interests.

With the growing emphasis on research output and scholarly publication, coupled with increasing competition for academic recognition and commercial applicability, researchers employ diverse linguistic strategies to enhance the perceived importance and impact of their studies. N. Millar et al. define the practice of authors employing exaggerated or evaluative language to embellish, promote, or amplify their research as 'rhetorical hype' [Millar et al. 2019].

By transforming empirical findings into disciplinary knowledge, academic discourse functions as a conduit for dialogue between scholars and their research communities. Consequently, the writing of a scientific paper is a socially constructed process through which authors textualize their research. Rather than merely reporting their findings, scholars strategically employ linguistic resources to justify their claims, negotiate perspectives, and foster alignment with their readership [Jiang 2019]. In essence, to gain peer recognition, researchers actively persuade their audience of the validity of their assertions. Such persuasion is not solely dependent on the reliability and validity of scientific experimentation but also on the deployment of rhetorical strategies [Jiang 2016; Liu et al. 2021]. Existing literature has identified multiple rhetorical devices used in academic writing, including metadiscursive markers such as **maximizers** (e.g., 'fully', 'totally', 'entirely') and **certainty stance adverbs** (e.g., 'actually', 'clearly', 'fascinating', 'important', 'remarkable') [Hyland, 2005]. Building on these findings, N. Millar et al. classify rhetorical hype as the strategic use of affirmative or emotionally charged language to promote academic research, often incorporating hyperbolic embellishment [Millar et al. 2019]. On one hand, rhetorical hype encompasses statements reflecting cognitive judgments or affective attitudes, such as 'The results **clearly** demonstrate that ...', as well as the use of evaluative language to magnify the novelty of a study, exemplified by

phrases like ‘The experiment used a **novel** technique...’ On the other hand, rhetorical hype extends beyond findings and research significance to also influence the portrayal of methodologies, as seen in expressions like ‘**normative** prescriptions for teaching design’.

Nevertheless, excessive promotional rhetoric in academic writing has sparked ethical concerns within the scientific community. Z. Master and D. B. Resnik (2013) and E.H. Ecklund et al. (2015) argue that scientific papers often exaggerate the validity of research methods and conclusions, thereby allowing subjective interpretations to overshadow objective scientific reasoning and compromising the integrity of scientific argumentation [Millar et al. 2020]. Similarly, T. Caulfield and C. Condit (2012) and T. Caulfield (2018) suggest that scholars resort to promotional rhetoric due to the pressure of academic publishing and the competitive commercialization of research. Studies by C.H. Vinkers et al. (2015) indicate a significant increase in the use of positively charged terms, such as ‘crucial’, ‘novel’, ‘innovative’ and ‘important, in medical literature over the past four decades, aiming to underscore the value of research findings and their clinical applications. Furthermore, longitudinal analyses by K. Hyland and F. Jiang (2021) demonstrate a rising trend in rhetorical hype across disciplines, reflecting a broader effort by authors to promote their research and attract readers and potential stakeholders in an era of heightened knowledge competition. Despite the growing body of research on the intersection of rhetorical hype and academic discourse, there is a lack of studies examining how Chinese learners of English utilize rhetorical hype in their academic writing, particularly in comparison with international journal publications. This issue will be explored in detail in Chapter 3.2.

### **1.4.3. Hedging strategy**

Hedging functions as a linguistic strategy enabling speakers and writers to present their assertions with a certain level of uncertainty or adaptability

[Alghazo et al. 2021; Kashiha, 2024]. Through the use of hedging devices, individuals can moderate the strength of their statements, avoiding excessive directness and positioning claims as perspectives rather than absolute facts [Hyland 2005]. The theoretical foundation of hedging can be traced back to G. Lakoff's seminal work (1973), in which he conceptualized it as a means of introducing ambiguity, caution, and nuance into communication. Within academic discourse, hedging not only dictates the extent to which an author commits to a claim but also plays vital pragmatic functions. These include fostering reader engagement, acknowledging subjectivity in interpretation, mitigating potentially face-threatening criticism [Chen & Zhang 2017], and maintaining a neutral or impartial stance [Yoon & Tabari 2023].

Moreover, K. Hyland (1996) underscored the significance of hedging within scholarly discourse, emphasizing its role in fostering uncertainty and facilitating academic dialogue. The perception of academic writing has shifted from being regarded as an objective and purely informational mode of expression to one that is inherently persuasive, where interaction and engagement with the audience are fundamental [Alghazo et al. 2021]. When strategically implemented, hedging helps authors navigate their stance on propositions while simultaneously negotiating power dynamics with their audience, including editors, reviewers, and readers [Adrian & Fajri 2023; Mu et al. 2015].

In addition, the proficient use of hedging reflects an author's pragmatic awareness, which pertains to their capacity to adjust linguistic choices based on varying situational contexts [Chen & Zhang 2017; Fraser 2010]. Such awareness ensures that academic compositions align with the communicative norms and expectations of distinct scholarly communities. Consequently, writers must possess the skill to deploy appropriate hedging mechanisms, as an absence or misuse of these structures can lead to miscommunication [Chen & Zhang 2017]. Despite its critical role in scholarly writing, research suggests that non-native English speakers – including those with advanced proficiency – often struggle to

implement hedging effectively [Almakrob 2020; Hyland & Milton 1997; Thabet 2018].

Extensive inquiries have investigated the ways in which speakers of different first languages employ hedging in academic English, spanning genres such as student dissertations and peer-reviewed journal articles (e.g., Adrian & Fajri, 2023; Almakrob, 2020; Chen & Zhang, 2016; Shafqat et al., 2022; Yagız & Demir, 2014). A significant portion of this research compares hedging tendencies between native and non-native English users. For example, A.Y. Almakrob (2020) examined hedging in doctoral dissertations written by Arabic-speaking scholars, while A. Shafqat et al. (2022) explored its use among Pakistani PhD candidates. Both studies found that non-native students incorporated hedges less frequently in their theses. Regarding published research, O. Yagız and C. Demir (2014) analyzed hedging strategies in ELT journals and discovered that Turkish scholars utilized hedging structures more frequently than native English writers. Similarly, C. Chen and L.J. Zhang (2016) observed that Anglophone researchers in applied linguistics incorporated a greater number of hedges in their conclusion sections than their Chinese counterparts. Additionally, D. Adrian and M.S.A. Fajri (2023) reported that Indonesian scholars working in soft sciences tended to use fewer hedging devices in their research compared to authors publishing in prestigious international outlets. Collectively, these findings suggest that native English speakers generally hedge their claims more often than non-native academics in research writing.

Prior investigations into hedging have largely centered on contrasts between native and non-native speakers. However, E. Smirnova and S. Strinyuk (2020) argue that the hallmark of an exemplary academic text is not necessarily tied to the author's linguistic background but rather to its acceptance in high-impact, peer-reviewed publications. Comparing learner-produced writing with professional scholarship is a justified and necessary endeavor, as expert texts have long been utilized as benchmarks for second-language (L2) academic

writing instruction and models for EAP learners. Additionally, although some studies have examined the hedging strategies employed by novice versus experienced writers, this line of research remains relatively underexplored. To date, two notable studies have addressed this gap. X. Gu and Z. Xu (2021) investigated the deployment of hedging in the results and discussion sections of unpublished engineering research papers, contrasting student compositions with those of seasoned scholars through the lens of metadiscourse. Similarly, E. Smirnova and S. Strinyuk (2020) examined hedging patterns in the business and management writing of Russian EFL learners in comparison to professional authors.

Furthermore, most existing studies have adopted corpus-based methodologies, often concentrating on hedging frequency without delving into the broader linguistic patterns associated with hedging expressions. While corpus-driven research on the collocational tendencies of hedging verbs remains relatively scarce, the role of hedging in journal articles extends beyond simple word choice. Authors deliberately employ linguistic resources to structure discourse, evaluate arguments, and engage their readers in a collaborative process of meaning-making. The co-occurrence patterns of hedging verbs function as explicit textual markers that facilitate these communicative objectives [Hyland 1998; Wang & Zeng 2021]. Addressing this research gap, this dissertation aims to investigate the collocational behavior of hedging verbs within academic texts written by Chinese MA students in linguistics, comparing their usage to that of expert scholars in the field, with the goal of elucidating the functional roles of hedging in academic writing. The detailed discussion will be demonstrated in Chapter 3.3.

#### **1.4.4. Authorial identity construction**

The intricate relationship between language and identity has long been a pivotal topic in sociolinguistics, with recent scholarship increasingly



emphasizing identity as an ongoing, dynamic process constructed through social interaction [Benwell & Stokoe 2006]. Rather than a fixed trait, identity is conceptualized as an individual's engagement with their social world, emerging through reciprocal exchanges with others. Language plays a central role in shaping and expressing one's sense of self, embedding individuals within networks of shared meanings, affiliations, and perspectives. Identity formation is a fluid process, continuously influenced by contextual factors, interpersonal relationships, and rhetorical choices in communication. Through language, individuals can assert their belonging to specific communities, differentiate themselves from others, or challenge group affiliations. This negotiation of identity is particularly significant in academic settings, where researchers establish professional credibility and shape their reputations within disciplinary discourses.

Navigating the expectations of academic writing while maintaining a distinct authorial presence requires skillful negotiation of disciplinary norms. This challenge is especially pronounced for EFL learners, who must balance adherence to standardized conventions with the need to construct a credible scholarly identity. Understanding identity as something performed through discourse necessitates a methodological approach that examines how individuals linguistically mark their presence in academic texts.

A scholarly work's credibility and reception rely not only on its methodological rigor but also on the effectiveness of its argumentation, often referred to as academic persuasion. For arguments to be compelling, they must align with the established conventions of academic discourse to resonate with readers [Jiang 2016]. Persuasion in academic writing is particularly evident when authors explicitly intervene in the discourse, organize their ideas, evaluate claims, and guide the audience in interpreting their arguments [Jiang 2019]. For instance, the statement "we show that" differs from "the results show that" in that the former foregrounds the author's active role in knowledge production, thereby

reinforcing their scholarly contribution. In this way, self-mentioning strategies serve as essential tools for constructing an authoritative presence and enhancing academic persuasiveness.

Linguistic elements such as personal pronouns (e.g., ‘I,’ ‘my,’ ‘we’) function as explicit markers of authorial presence, fostering engagement between writers and readers in academic texts. These **self-mentions** are crucial for asserting a writer’s stance and reinforcing the persuasive power of their arguments. Despite their significance in knowledge dissemination, novice Chinese researchers often struggle with the effective integration of self-mentions in their academic writing, which can hinder their ability to construct coherent arguments and establish a strong scholarly identity [Wang & Jiang, 2019]. The patterns of self-mention usage among Chinese novice researchers and expert academics will be examined in detail in Chapter 3.4.

## CHAPTER I CONCLUDING REMARKS

We have conducted a comprehensive review of the definitions, categorizations and research perspectives concerning metadiscourse. In this part, we intend to summarize the consensus and remaining controversies regarding the research on metadiscourse.

1. Converging understanding: Consensus on metadiscourse in scholarly research

Presently, the understanding of metadiscourse among scholars has converged on the following points: (1) Regardless of the nomenclature used, metadiscourse is universally recognized as a linguistic phenomenon. (2) Metadiscourse constitutes a fuzzy category, lacking clear boundaries between metadiscourse and non-metadiscourse, necessitating consideration of various criteria for its delineation. (3) Metadiscourse resources form a continuum, with the categorization of its types not being absolute but instead a matter of extent. (4) Metadiscourse studies are examined from various perspectives, including logical philosophy, semiotics, pragmatics, communication studies, and rhetoric. (5) Metadiscourse projects exhibit multifunctionality, capable of simultaneously serving multiple functions within the same discourse context. (6) Metadiscourse is context-dependent, requiring the inclusion of contextual factors in defining it and determining its functions. (7) Scholars have generally agreed on the distinction between ‘metatext’ and ‘metadiscourse’: ‘metatext’ refers to the more limited concept that includes only textual elements, whereas ‘metadiscourse’ is a more comprehensive term that contains both interpersonal and textual elements. The consensus reached by scholars from different fields regarding metadiscourse research enhances our comprehension of the essence of metadiscourse from a multidisciplinary perspective. A. Crismore (1989) has pointed out that such an approach can more precisely define this concept, providing a more comprehensive classification for subsequent discourse studies. For this reason,

this thesis adopts this broader term ‘metadiscourse’ and employ its ‘interpersonal model’ raised by K. Hyland in this paper [Hyland 2005:49].

2. Persisting disputes: Metadiscourse controversies in scholarly research discourse

(1) The debate concerning propositions and non-propositions. This issue revolves around the nature of metadiscourse and whether it influences the propositional content of discourse. Metadiscourse phenomena, since their inception, have frequently been labeled as ‘non-propositional content’, with proponents asserting its independence from truth conditions, involving only textual and interpersonal functions, and lacking conceptual functions or adding any propositional information. Sometimes, the importance of metadiscourse in communication even surpasses that of the topic. The controversy over whether metadiscourse constitutes propositions essentially revolves around whether truth conditions should be used as the criteria for evaluating metadiscourse. In conventional contexts, readers primarily focus on the suitability of the interpersonal meanings revealed by propositions and the logical coherence of discourse, rather than their correspondence with truth values in the world. Metadiscourse involves readers in the discourse process, capturing the author’s intentions and guiding the author in producing coherent, cohesive, and reader-friendly discourse, thereby facilitating the understanding of the basic discourse. Therefore, the focus of metadiscourse research should be on the interactive or interpersonal aspects of metadiscourse markers, rather than discussing their truth values.

(2) The controversy concerning the relationship between metadiscourse and base discourse. Many researchers view metadiscourse as auxiliary discourse that helps authors organize and convey base discourse. This categorization of base and auxiliary discourse relegates metadiscourse to a subordinate position within discourse. This has elicited objections from some scholars. K. Hyland (2005) argues that metadiscourse not only supports propositional content but also

contributes to the coherence and persuasiveness of propositional content. Propositional discourse and metadiscourse are equal; they should be combined to constitute the two aspects of the entire discourse, directing attention to two primary entities: entities in the world and entities in the discourse. R. Mao (1993) attempts to blur the distinction between base and auxiliary discourse. He believes that from an etymological perspective, the prefix ‘meta-’ in this term refers to ‘fundamental issues’, and metadiscourse deals with fundamental communicative issues, namely, how to establish interpersonal connections and maintain relationships between discourses. Therefore, metadiscourse is more fundamental than so-called basic discourse and is closer to the overall goal of language use.

The consensus reached by scholars regarding metadiscourse research demonstrates the universality and breadth of this linguistic phenomenon, representing the concerted efforts of scholars in theory and practice and laying the groundwork for further research and communication. The controversies surrounding current metadiscourse research also underscore the complexity of this phenomenon, serving as the impetus for scholars to deepen their research and broaden their perspectives.

3. Metadiscourse and rhetorical strategies: Enhancing persuasion through rhetorical hyping strategy, hedging strategy and authorial identity construction.

Metadiscourse, a strategic linguistic tool, is central to communication, allowing authors to embed their perspectives, evaluations, and interpretative judgments into texts while anticipating audience reception. It structures ideas, enhances engagement, ensures coherence, conveys epistemic certainty, and steers comprehension, serving persuasive functions. Rooted in ancient Greek rhetoric, metadiscourse supports these modes of persuasion: *logos* (logical coherence through explicit argument connections), *ethos* (credibility via the author’s expertise and authority), and *pathos* (emotional engagement by aligning with audience expectations). This section analyzes the interplay between metadiscourse and these modes of persuasion, focusing on three strategies:

rhetorical hyping (amplifying claims to emphasize significance), hedging (mitigating assertions to express caution or uncertainty), and authorial identity construction (crafting a credible persona to bolster trust). These strategies exemplify how metadiscourse operates as a dynamic mechanism for persuasion, balancing logical rigor, ethical credibility, and emotional resonance. The discussion clarifies their roles within the dissertation's broader exploration of rhetorical tactics, underscoring metadiscourse's multifaceted capacity to shape discourse and influence audience interpretation.

## CHAPTER II. ENGLISH FOR ACADEMIC PURPOSES

This chapter consists of definitions, characteristics, and methodologies commonly employed in English for academic purposes (EAP) studies. Furthermore, concluding remarks of this chapter will be demonstrated in the end.

### 2.1. Definitions of EAP

EAP, in a broad sense, refers to the English language as a medium for academic studies and communication in the academic discourse. Based on the varying types of source texts, academic English can be categorized into two distinct branches: English for General Academic Purposes (EGAP) and English for Specific Academic Purposes (ESAP) [Jordan 1997]. EGAP concentrates on the common aspects of discipline-specific English and aims to develop students' oral and written communication skills for professional studies, such as lecture-listening, note-taking, literature review, academic report, and academic writing. ESAP emphasizes on academic language learning in specific specialties (e.g., Linguistics, Sociology, Biology and Chemistry), which involves academic vocabulary, academic grammar, academic genre, etc.

EAP varies in their definitions. K. Hyland and L. Hamp-Lyons define it as «language research and instruction that focuses on the specific communicative needs and practices of particular groups in academic contexts. It means grounding instruction in an understanding of the cognitive, social and linguistic demands of specific academic disciplines. This takes practitioners beyond preparing learners for study in English to developing new kinds of literacy: equipping students with the communicative skills to participate in particular academic and cultural contexts» [Hyland, Hamp-Lyons 2002: 2]. R. Scarcella defines EAP as «a variety or a register of English used in professional books and characterized by the specific linguistic features associated with academic disciplines» [Scarcella 2003:

9]. In other words, EAP covers abstract-reading, lecture note-taking, literature reviewing, bibliographic notes, research reports, case studies, and research projects. This definition relates to the writing system of EAP and its unique academic norms and emphasizes the disciplinary characteristics. Nevertheless, N. X. Wei argues that academic English is essentially a subcategory of English, a variant of English used by academic experts for knowledge construction, information exchange and dissemination of professional academic knowledge to advance their disciplines, whose discourse participants are academics with professional disciplinary knowledge [Wei 2016]. The description of discourse acts involves academic vocabulary, grammar, and their semantic meaning and pragmatic function, and non-discourse acts involve professional formulae, diagrams, and other metalanguage, and the discourse effect involves the dissemination of professional knowledge and the exchange of information. In addition, F. Jiang argues pedagogically and scholarly that EAP takes the EFL learners' or experts' text of the bases and describes or compares the linguistic characteristics, genres and discourse practices shared by the target discourse community through empirical evidence. The aim is to raise students' awareness of genre and rhetoric, to make them understand the purpose of academic communication and the culture of the discipline, so that they can better integrate into the discourse community of the discipline to which they belong and construct their own academic identity [Jiang 2019; Lou 2022]. This definition clearly presents the research purpose of EAP and highlights the essential feature that academic English serves foreign language teaching.

EAP can provide learners with 'textual norms' in a specialty field or disciplinary field on the basis of textual analysis of academic writing [Hyland 2018], and has the characteristics of being well-defined, regular and teachable [Scarcella 2003], and thus closely linked to foreign language teaching and research. Unlike language research in other genres, EAP research is closely integrated with teaching and learning, and the rationale for the research is to



provide guidance and feedback on classroom teaching practices or academic English writing by describing the linguistic features of EAP.

## 2.2. Characteristics of English for academic purposes

The main purpose of academic English research is to serve English academic writing of teacher or researchers in colleges and universities, and the linguistic characteristics of academic English texts are the basic starting point for academic English teaching.

C.E. Snow and P. Ucceli concluded the linguistic features of EAP mainly into the following 8 aspects in Table 1 [C. E. Snow & P. Ucceli 2009: 118–121].

**Table 1 The linguistic features of EAP**

- 
1. Interpersonal stance: detached/distanced and authoritative stance
  2. Information load: conciseness and density
  3. Organization of information: constituency; explicit awareness of organized discourse; autonomous text; stepwise logical argumentation/unfolding, tightly constructed
  4. Lexical choices: high lexical diversity; formal/prestigious expressions; precision and abstract/technical concepts
  5. Representational congruence: complex/congruent grammar; compact grammar; abstract concepts as agents
  6. Genre mastery: school-based genres; discipline-specific specialized genres
  7. Reasoning strategies: specific reasoning moves valued at school; Discipline-specific reasoning moves
  8. Disciplinary knowledge: abstract groupings and relations; disciplinary taxonomies and salient relations; knowledge as constructed

All of these linguistic features must be matched with the following three cognitive skills: (1) Genre mastery: Academic English writing skills encompass narration, description, explanation, etc., with educational contexts also necessitating the composition of laboratory reports and persuasive essays; (2) Reasoning strategies: Academic English demands not only basic argumentation and persuasion skills but also requires discipline-specific reasoning steps; (3) Disciplinary knowledge: Academic writing necessitates comprehension, categorization, and mastery of the inherent connections among terms; not solely the mastery of factual knowledge but also the construction of professional knowledge. As proficiency in these three skill areas increases, academic English writers gradually learn to utilize the features of academic language to enhance these skills, specifically mastery of genres.

As evident, there are notable distinctions between academic English and everyday English, and these variances can pose significant challenges for learners. Therefore, pinpointing the disparities between academic English and daily English aids learners in academic writing. K. Hyland delineates eight features that set academic English apart from daily English:

(1) Be more explicit about its structure and purpose: This mirrors the highly structured and conventionalized aspects of EAP, where a clear structure can mitigate readers' comprehension difficulties, bolster textual persuasiveness, and facilitate academic communication objectives.

(2) Use more citations to support arguments: Academic writing necessitates ample citations to substantiate arguments, which not only underscore the disciplinary nature of academic research but also persuade readers to endorse the author's standpoint.

(3) Focus on actions rather than actors: A typical manifestation of this is the prevalence of nominalization and passive voice constructions in academic texts.

(4) Limit the use of rhetorical questions compared to what students typically include in school essays: Academic texts predominantly feature declarative sentences, and rhetorical questions are used fewer, except when articulating research inquiries.

(5) Be generally intolerant of digressions: This entails that the argumentation in academic texts revolves around key concepts, without introducing irrelevant content.

(6) Be cautious in making claims: Authors of academic texts must adopt a cautious stance, furnish ample and reasoned arguments, and refrain from employing overly definitive language, allowing room for negotiation.

(7) Package processes as things: Academic texts tend to employ noun structures over verb structures when delineating events.

(8) Spell out steps in an argument and connections between sentences very clearly: The logical connections between sentences in academic texts are explicit, and their relationships are reinforced through grammatical cohesion (e.g., conjunctions) and lexical cohesion (e.g., collocations) [Hyland 2017: 26].

Studies of EAP typically investigate the linguistic mechanisms, through which these features are manifested in academic discourse. Methodologies predominantly employed include genre analysis, discourse analysis, and corpus analysis, all of which are grounded in authentic academic texts. Early genre analysis and discourse analysis techniques are non-corpus approaches that do not necessitate the use of computerized statistical software. In contrast, corpus methodologies entail studies that utilize corpus software along with corpus linguistic techniques. Each of these three methodologies will be expounded upon in the subsequent section.

## **2.3. Methodologies in English for academic purposes**

This section will examine three predominant methodologies commonly utilized in EAP studies, namely genre analysis, discourse analysis, and corpus approach.

### **2.3.1. Genre analysis**

Early investigations into EAP often relied on a corpus of texts from the same genre, primarily concentrating on the organizational and representational aspects of the texts. J. Swales defines genre as a category of communicative events expressed in various languages or language variants within social activities, serving specific purposes and agendas. To accomplish communicative goals, individuals within a particular community adhere to shared genre conventions and employ similar content and structure in their discourse. The efficacy of individuals' communicative actions partially hinges on their genre awareness. Genre constitutes a communicative event characterized by a distinctive internal structure and a high level of consensus. Writers within the same discourse community must comprehend the structural and linguistic features as well as the conventions of a particular discourse genre [Swales 1990].

Research articles serve as the primary genres for the global dissemination of expertise and scholarly accomplishments, being the primary genre selected for EAP studies. A genre-based approach can effectively delineate the structural, linguistic, and rhetorical characteristics of one genre, aiding learners in understanding the requirements of academic writing and furnishing a structured framework for discourse analysis. This ensures that classroom instruction aligns with learners' needs, thereby offering an effective method for EAP teaching.

Genre analysis entails the examination of conventional linguistic behaviours within academic or professional contexts, with different theoretical

schools emphasizing various aspects, whether focusing on the typicality of behaviour, regularity of goal-oriented social processes, or the consistency of communicative purposes. What they share in common is, firstly, that genre serves as a reflection of disciplinary culture and discourse reality. Secondly, genre centres on the conventionalized communicative events within a disciplinary field or professional practice. Thirdly, all disciplines or professional genres exhibit their own coherence, often referencing internal (textual and discursive) or external (contextual and disciplinary) factors. Fourthly, genre constitutes an identifiable communicative occurrence defined by a collection of shared communicative purposes understood by members of a professional or academic community where such events occur. Fifthly, genres are constructions with specific structures and traditions, distinguishable based on intended purposes and shared grammatical resources. Sixthly, members of a professional group possess more extensive knowledge and comprehension of professional academic practices compared to novices, newcomers, or outsiders.

The constrained nature of genres delineates the distinctions among genres and various modes of analysis. A prevalent rhetorical pattern observed in technical academic papers is the ‘problem-solution’ pattern [Flowerdew 2003]. M. Hoey (2001) further segmented it into ‘Situation-Problem-Response-Evaluation/Result’. This model can be employed to assess whether the writers’ textual strategy facilitates more effective communication with the reader.

In accordance with the overarching discourse structure of academic papers, they typically comprise four moves, namely introduction, method, results, and discussion, collectively known as IMRD moves. Each discourse move possesses its own structural characteristics, and distinct discourse functions, thereby necessitating diverse linguistic mechanisms to fulfill these functions [Swales 1990].

J. Swales proposed the ‘Creating a Research Space’ analysis model for the structure of **introductions** in academic papers, which is applicable to the

structural analysis of introductions in various scientific papers [Swales 1990]. The structural analysis is delineated into three moves: Move 1 begins with the establishment of the research field, involving the research topic, the generalization of the theme, and the review of the research before. Move 2 proceeds with the identification of the research limitations, which includes countering existing arguments, pinpointing research gaps, highlighting current issues, and extending upon previous research outcomes; Move 3 elucidates how the research group can address the shortcomings of the research area, comprising stating the research purpose, describing the current progress of relevant research, and introducing the paper's structure. J. Swale's motivation in proposing this model was to standardize the 'Introduction' in academic papers as a genre agreed upon by academia. He believed that the purpose of an academic paper was to disseminate one's scientific results to peers. Genre analysis should not only focus on describing linguistic features or style but also consider communicative purposes, as each move is designed with a specific communicative intent.

Similarly, the genre characteristics of **abstract** in academic discourse have been explored by analyzing the moves within abstracts. The representative five-move model, consisting of introduction, aim, method, results, and conclusion, is employed to delineate the research scope, describe the study's objectives, outline the research methodology, report the findings, and summarize the main conclusions.

In addition, J. Swales explored the structural features of academic papers, emphasizing the depiction of the framework for different genres. He focuses on the predictable formulaic aspects of the different moves by presenting the structure of the moves in the paper and their typical linguistic realizations in the text, arguing that a large number of textual observations can summarize the structure of the academic genre and the features of its language use, which have regularity and repeatability, and that such regular linguistic features allow academic communication in different research fields to be carried out effectively.

Research papers are usually composed of four discourse moves, namely **introduction**, **method**, **result**, **discussion** and **conclusion**, each of which has its own structural features, and different discourse moves have different functions in the discourse, thus requiring different linguistic means to achieve these functions. J. Swales' findings provided a suitable framework for later studies of EAP writing and teaching, and its framework was particularly suited to analyzing features of textual superficial structure and the relationship between these features and the discipline.

E. Zanina underscores the pivotal role of move-based genre analysis in crosslinguistic academic communication research and EAP/ESP pedagogy. However, she identifies a gap in linguistic or pedagogically motivated studies comparing English and Russian research articles. Employing K. Hyland's 5-move model, her study delves into cross-linguistic disparities in abstract move structures within research articles on management. Analyzing 20 unstructured RA abstracts in both languages, E. Zanina observes that while English abstracts largely conform to K. Hyland's model, Russian abstracts predominantly exhibit a three-move structure. Notable discrepancies include English authors' detailed method and result descriptions compared to Russian writers' brevity. Despite limitations in sample size and descriptive nature, E. Zanina's research holds potential interest for genre analysts and L2 theorists [Zanina 2017].

T. Dudley-Evans was among the pioneering scholars to acknowledge the significance of comprehending the social and disciplinary contexts of academic discourse, and he proposed that academic discourse exhibits disciplinary characteristics and traditions. He contends that students' awareness of these characteristics facilitates the EAP writing process and underscores the importance of close collaboration between English instructors and instructors of specialized courses [Dudley-Evans 1994]. T. Dudley-Evans builds upon J. Swales' examination of discourse moves in research papers and directs attention

to other genres, wherein the structures and linguistic patterns of various moves are frequently regular and recurrent.

K. Hyland's research has also made a significant contribution to EAP writing and teaching, with a focus on examining how language is utilized to construct academic knowledge, foster social relationships, and convey disciplinary attributes. His research spans a broad spectrum of topics, including comparative analyses of graduate dissertations across disciplines, assessment, exemplification, and reconstruction in dissertations, and providing feedback on learners' dissertations [Hyland 1998, 2004, 2005, 2017].

These studies concentrated on disciplinary disparities in academic paper writing and revealed that different disciplines tend to employ distinct structures and linguistic expressions in various discourse moves (e.g., introduction, results, and abstract). These findings were corroborated in subsequent corpus-based empirical studies.

### **2.3.2. Academic discourse analysis**

Academic discourse analysis (ADA) has emerged as a pivotal methodology in EAP research, particularly for investigating metadiscourse features that mediate writer-reader interactions [Hyland 2005]. Rooted in systemic-functional linguistics and genre analysis, ADA examines how language constructs academic knowledge, emphasizing rhetorical strategies, coherence-building, and stance-taking [Swales 1990]. Researchers employ ADA to deconstruct disciplinary conventions, revealing how writers employ metadiscourse – such as hedges, boosters, and transitions – to guide readers through arguments [Hyland & Tse 2004]. This approach aligns with the comparative focus of this study, as it enables systematic contrasts between Chinese EFL learners and expert writers in terms of their metadiscursive choices. By foregrounding context-specific language use,



ADA provides a framework for analyzing how cultural, linguistic, and educational backgrounds shape academic writing practices.

A central tenet of ADA is its emphasis on the social function of metadiscourse, which positions academic writing as a dialogic act [Crismore et al. 1993]. Studies applying ADA to EFL contexts have demonstrated that learners often underuse or overuse interactive markers (e.g., frame markers, code glosses) compared to expert norms, reflecting differing assumptions about audience engagement [Ädel 2006]. For instance, Chinese EFL learners' reliance on overt textual organizers may stem from L1 transfer or pedagogical traditions prioritizing structural clarity over reader-oriented persuasion [Hu & Cao 2015]. Conversely, expert writers deploy metadiscourse strategically to balance authority with collegiality, a skill honed through disciplinary enculturation [Dafouz-Milne 2008]. ADA's analytical lens thus illuminates how metadiscursive competence evolves across proficiency levels and cultural contexts, offering insights into the challenges faced by EFL learners in approximating expert conventions.

Critically, ADA underscores the role of disciplinary and cultural specificity in shaping metadiscourse patterns. Comparative studies reveal that Chinese academic writing, influenced by Confucian rhetoric, often prioritizes collective voice and indirect argumentation, contrasting with Anglophone norms of explicit self-mention and critical engagement [Connor 1996; Mauranen 1993; N]. Such cultural divergences complicate EFL learners' adoption of metadiscourse strategies deemed effective in Western academia. For example, Chinese learners may avoid boosters to maintain humility, inadvertently weakening persuasive impact [Lee & Casal 2014]. ADA methodologies, which integrate corpus-based analyses with ethnographic insights, enable researchers to disentangle these cultural-linguistic intersections, highlighting tensions between L1 rhetorical traditions and L2 academic expectations [Hyland 2017].

Recent advancements in ADA have leveraged computational tools to enhance the granularity of metadiscourse analysis. Corpus-driven studies, utilizing platforms like AntConc or Sketch Engine, quantify metadiscursive frequency and distribution across learner and expert corpora, revealing statistically significant variations [Biber et al. 1999]. For instance, L. Anthony (2019) employed keyword analysis to identify overused transitionals in EFL writing, linking these patterns to L1 interference. However, critics caution against over-reliance on automated tools, advocating for mixed-methods approaches that incorporate qualitative text analysis to capture pragmatic nuances [Flowerdew 2015]. Such methodological rigor is vital for this study, which combines quantitative corpus comparisons with discourse-based interviews to contextualize metadiscursive deviations.

Ultimately, ADA's interdisciplinary orientation bridges linguistic theory and pedagogical practice, offering actionable insights for EAP instruction. By diagnosing gaps in learners' metadiscourse repertoires, ADA-informed research can tailor interventions that address both linguistic accuracy and rhetorical appropriateness [Hyland 2005]. For Chinese EFL learners, this may involve explicit training in audience adaptation, hedging strategies, and disciplinary-specific conventions – areas where expert writing provides benchmark models. As this study applies ADA to a cross-cultural comparative framework, it contributes to ongoing debates about global academic literacy standards while advocating for culturally responsive pedagogies that honor learners' multilingual repertoires [Canagarajah 2013].

### **2.3.3. Corpus approach**

Corpus-based EAP studies emerged in the 1980s and typically do not focus on overall discourse and its social context, often termed as a 'bottom-up' approach. This approach involves examining large volumes of textual data to

gather frequency information and distribution of surface features in language use. T. Johns conducted pioneering research in discipline-specific areas and demonstrated the potential of corpus and concordance techniques for language analysis and teaching practices [Johns 1986].

Corpus approaches have played a vital role in delineating the linguistic characteristics found in academic discourse. Research at the lexical level suggests that vocabulary teaching in EAP classrooms should encompass three categories of English vocabulary: high-frequency vocabulary, academic vocabulary, and specialized technical vocabulary [Coxhead, Nation 2001]. High-frequency vocabulary, forming the foundation of language use, comprises approximately 2000-word families and typically constitutes 80% of the total words in academic texts. Academic vocabulary consists of around 570 words [Coxhead 1998], occurring frequently in academic discourse but less so in other contexts. Academic vocabulary is crucial for EAP learners, accounting for about 10% of total words in academic texts. Unlike the first two groups, technical vocabulary is discipline-specific and varies across disciplines, constituting about 5% of total words in academic texts.

This differentiation aids in establishing vocabulary teaching objectives for EAP courses, particularly for discipline-specific ones. As EAP learners already possess some general vocabulary, research is needed to identify words that are not part of general vocabulary but appear frequently across disciplines.

The corpus approach has highlighted the phenomenon of lexical bundles in discourse, considered crucial for distinguishing disciplinary characteristics in written texts. Effective writing entails an awareness of the implications of lexical expressions and mastering the appropriate usage of lexical bundles. Studies based on EAP corpora of native speakers have further validated the existence of distinctive terms in EAP writing by examining recurrent words and word combinations [Biber, Johansson, Leech, Finegan 1999; Biber 2006], as well as chunks [Cortes 2004; Hyland 2008]. These studies typically explore the usage

and functions of discourse bundles in different contexts. While their impact on syllabus and material design remains limited, they demonstrate that a corpus perspective can inform the design of effective inductive and deductive exercises to help students master high-frequency EAP words and word combinations.

Corpus-based EAP research also encompasses discourse analysis and move analysis, focusing on describing and explaining text properties. These studies suggest that corpus-based methods can be utilized to examine language variation.

EAP studies focus on portraying and interpreting specific linguistic forms or structures as they appear in discourse, or on discourse features embodied in different genres. Both corpus and non-corpus approaches highlight substantial differences between disciplines in genre choice and genre structure. They examine stylistic differences, syntactic characteristics, and rhetorical functions of different discourse moves within academic discourse. Corpus-based research offers a more reliable quantitative investigation, based on analysis of large real language data sets using scientific methods to present reliable data. However, non-corpus approaches have developed theoretical frameworks allowing for a broader dimension of EAP research, contributing to the improvement of EAP teaching and writing, particularly in analyzing organization of genre-specific and discipline-specific academic discourse and typical linguistic patterns.

#### **2.3.4. Critical discourse analysis**

Critical discourse analysis (CDA) provides a robust framework for interrogating how metadiscourse practices in academic writing reflect power relations, cultural ideologies, and identity negotiation [Fairclough 1995; Hyland 2005]. Rooted in the premise that language is a social practice, CDA transcends surface-level linguistic analysis to expose how metadiscursive markers – hedges, boosters, transitions, and self-mentions – mediate authority, credibility, and

disciplinary alignment [Hyland 2005; Van Dijk 2008]. In EAP research, CDA has been instrumental in deconstructing disparities between novice and expert writers, particularly in how learners navigate Anglophone conventions while negotiating L1 rhetorical traditions [Canagarajah 2002]. For instance, studies reveal that Chinese EFL learners often underuse interactional metadiscourse (e.g., “we propose”) due to cultural preferences for indirectness, contrasting with Anglophone experts’ strategic deployment of such features to assert argumentative clarity [Li & Wharton 2012]. CDA thus positions metadiscourse as a site of ideological tension, where writers balance disciplinary expectations with sociocultural identities.

CDA’s three-dimensional model – text, discourse practice, and sociocultural context [Fairclough 1995] – aligns with comparative metadiscourse studies. Textual analysis identifies quantitative patterns (e.g., frequency of hedges in learner vs. expert corpora), while discourse practice examines how these choices reflect socialization into academic communities [Curry & Lillis 2017]. For example, Chinese learners’ limited use of self-citation may signal unfamiliarity with Western norms of self-promotion, whereas experts’ strategic citations reinforce epistemic authority [Flowerdew 2008]. At the sociocultural level, CDA interrogates how global academic hierarchies privilege Anglophone metadiscourse, marginalizing alternatives [Curry & Lillis 2017]. This multilayered approach enables researchers to trace metadiscursive differences not merely as linguistic deficits but as manifestations of unequal access to dominant academic repertoires.

CDA’s integration with corpus linguistics further strengthens its applicability to metadiscourse research. Tools like AntConc or Sketch Engine allow systematic identification of metadiscursive patterns across learner and expert corpora, while CDA contextualizes these patterns within power dynamics [Baker et al. 2008]. For instance, corpus-driven studies reveal that Chinese learners overuse formulaic transitions (e.g., “furthermore”), reflecting

pedagogical emphases on structural rigidity, whereas experts employ varied metadiscourse to engage readers dialogically [Hyland 2005]. However, CDA cautions against decontextualized frequency counts, advocating for qualitative insights into how learners experience metadiscourse – such as anxiety over appropriate self-mention – in cross-cultural contexts [Ivanič 1998]. This duality ensures methodological rigor while centering writers’ agency and lived realities.

Critics argue that CDA risks overemphasizing structural power imbalances, potentially framing learners as passive subjects of dominant discourses [Blommaert 2005]. Yet, recent EAP scholarship adopts a transformative CDA lens, exploring how learners reappropriate metadiscourse to hybridize L1/L2 practices. For example, Chinese EFL writers may blend Confucian humility with Western assertiveness via phrases like ‘this study tentatively suggests’. Such practices challenge deficit narratives, repositioning learners as strategic agents navigating “third spaces” [Canagarajah 2002]. CDA’s reflexivity – researchers acknowledging their own ideological stances – is vital here, ensuring analyses avoid ethnocentric biases [Wodak & Meyer 2016].

### **2.3.5. Thematic analysis**

Thematic analysis, rooted in Systemic Functional Linguistics (SFL), examines how writers organize information flow through theme (the starting point of a clause) and rheme (the remainder of the clause) to achieve coherence and cohesion [Halliday & Matthiessen 2014]. In EAP research, this approach has been pivotal in identifying how EFL learners structure arguments compared to expert writers. Studies reveal that Chinese EFL learners often struggle with thematic progression patterns, such as abrupt shifts to “brand new themes” or overusing “constant theme patterns,” leading to fragmented coherence. For instance, novice writers may reiterate the same theme without advancing arguments, whereas experts employ “zig-zag” or “multiple theme” patterns to guide readers through logical transitions [Fitriati & Gayatri 2021]. These

disparities highlight the role of cultural and pedagogical influences in shaping rhetorical strategies.

A key challenge in thematic analysis lies in addressing cross-linguistic differences. Chinese learners' thematic choices often reflect L1 rhetorical traditions, such as indirectness or preference for high-context communication, which clash with Anglophone norms favoring linear progression and explicit topic shifts [Gu et al. 2021]. For example, learners may prioritize Confucian ideals of humility by downplaying assertiveness, resulting in underdeveloped argumentation. Corpus-based studies comparing Chinese EFL essays and L1 expert writing further demonstrate that learners overuse textual themes (e.g., “however,” “furthermore”) as structural crutches, whereas experts deploy interpersonal themes (e.g., “we propose”) to engage readers dialogically [Ruan 2020]. Such findings underscore the interplay between linguistic competence and cultural identity in academic writing.

Pedagogically, thematic analysis informs interventions to enhance coherence. Experimental studies show that explicit instruction in thematic progression – such as teaching “reiteration” and “hypertheme” strategies – significantly improves EFL learners' ability to construct logically connected texts. For example, learners trained to use “split progression” (linking themes hierarchically) produced essays with higher coherence scores than those taught conventional grammar rules [Fitriati & Gayatri 2021]. However, critiques argue that rigid adherence to SFL frameworks risks oversimplifying the dynamic, context-dependent nature of writing [Gu et al. 2021]. Hybrid approaches, integrating corpus tools to visualize theme-rheme patterns, have emerged to balance structural rigor with flexibility [Zhang & Sheng 2023].

Recent advancements in digital pedagogy further expand thematic analysis's scope. MOOCs and AI-driven platforms now offer personalized feedback on thematic progression, enabling learners to compare their writing against expert corpora in real time. For instance, tools like Oxford WordSmith

analyze theme-rheme distributions in student drafts, highlighting overused transitions or disjointed arguments. Yet, such technologies must account for disciplinary variation: STEM writers favor “linear” progression, while humanities texts employ “spiral” thematic development to explore multifaceted arguments. Adapting these tools to discipline-specific needs remains critical for sustainable EAP pedagogy [Gu et al. 2021; Ruan 2020; Zhang & Sheng 2023].

Besides there are studies exploring how thematic progression intersects with identity construction in multilingual writing. Studies suggest that Chinese learners negotiating “third spaces” between L1 and L2 norms hybridize thematic patterns, blending indirectness with Western assertiveness. Additionally, corpus-assisted diachronic analyses could trace how learners’ thematic choices evolve with proficiency, offering insights into scaffolding strategies. As academic writing increasingly prioritizes global accessibility, thematic analysis remains indispensable for decoding – and bridging – rhetorical divides in cross-cultural scholarship [Fitriati & Gayatri 2021].



## CHAPTER II CONCLUDING REMARKS

EAP encompasses a broad spectrum of language study and teaching, focusing on the specific needs and services of individuals in academic contexts. Various definitions of EAP highlight its theoretical and pedagogical significance, ranging from its focus on specific linguistic features associated with academic disciplines to its role in facilitating foreign language teaching and research.

The characteristics of EAP, as outlined by scholars such as C.E. Snow and P. Uccelli, underscore the linguistic complexity and genre-specificity of academic English, necessitating proficiency in genre mastery, reasoning strategies, and disciplinary knowledge. These characteristics pose significant challenges for learners transitioning from daily English to academic discourse.

EAP methodologies encompass a range of analytical approaches that collectively explore the linguistic, structural, and discursive features of academic texts. Genre analysis examines the conventional structures and communicative moves that define academic writing, revealing how disciplinary cultures shape textual organization and rhetorical strategies. Academic discourse analysis delves into metadiscourse markers and writer–reader interactions, highlighting the social and cultural influences embedded in scholarly texts. Corpus approaches provide quantitative insights by analyzing large datasets to uncover patterns in vocabulary, lexical bundles, and syntactic structures, thereby informing effective pedagogical practices. Critical discourse analysis interrogates how power dynamics and cultural ideologies are reflected in metadiscourse choices, exposing the negotiation of authority and identity. Finally, thematic analysis, drawing on Systemic Functional Linguistics, explores the organization of information through theme and rheme, illustrating how thematic progression affects coherence and argumentation. Together, these methodologies offer complementary perspectives that not only deepen our understanding of academic

writing practices but also guide targeted instructional strategies for both novice and expert writers.

## CHAPTER III. THE USE OF RHETORICAL STRATEGIES OF METADISCOURSE IN THE ACADEMIC WRITING OF CHINESE EFL LEARNERS AND EXPERTS

Utilizing corpora that we constructed ourselves, the upcoming section will delve into an examination and comparison of metadiscourse characteristics. These features encompass certainty stance adverbs, maximizers, hedges, self-mentions, and engagement markers. We will explore their usage within three distinct rhetorical strategies: rhetorical hype, hedging strategies, and authorial identity construction in the academic texts produced by Chinese EFL learners and experts. Additionally, this study will investigate the underlying mechanisms that account for the observed differences between these two groups.

### 3.1. Research methodology

The dataset examined in this research consists of master's theses in English linguistics authored by Chinese students, alongside research articles contributed by established scholars in reputable academic journals.

To minimize **disciplinary variation**, this dissertation specifically focuses on the field of **English applied linguistics** as its primary research domain. Besides, English applied linguistics was selected as the focal discipline because most master's theses in this field in China are written in English and are publicly available through the CNKI (China National Knowledge Infrastructure) database. Furthermore, the selection and retrieval of MA applied linguistics theses strictly adhered to the **standardized academic structure** based on the IMRD (Introduction, Methods, Results, and Discussion) framework [Swales 1990]. This approach not only facilitated systematic data collection but also ensured the consistency of representative datasets that reflect the academic writing practices of Chinese learners.

By concentrating on applied linguistics in English, the study secures a sufficient corpus of scholarly essays, thereby enhancing both the practicality and reliability of the comparative analysis. A purposive sampling approach was applied, with predefined selection criteria guiding the inclusion of materials in the corpus.

To ensure high-quality research outputs, journal sources had to be indexed in Scopus and acknowledged as leading core publications in Applied linguistics field. The selection process did not hinge on whether authors were ‘native English speakers’ but rather on the Source Normalized Impact per Paper (SNIP) indicator. Because it accounts for disciplinary differences in citation practices, enabling fair comparisons across interdisciplinary subfields (e.g., TESOL, corpus linguistics) by weighting citations based on a field’s ‘citation potential’. SNIP also emphasizes citation quality over quantity (e.g., reducing bias from fields with longer reference lists), offering stable, field-normalized benchmarks (1.0 = median impact). Journals like *Journal of Second Language Writing* (SNIP: 2.838) reflect above-median influence in their niches, ensuring the corpus prioritizes impactful, contextually relevant research while balancing diverse subfields.

The dataset consists of two electronic corpora: (1) the Chinese Linguistic MA Theses Corpus (CLMA\_C), encompassing 50 master’s theses in applied linguistics from 36 Chinese universities, with a total token count of 804,025; and (2) the International Linguistic Journal Articles Corpus (ILJA\_C), comprising 100 published articles from influential applied linguistics journals, with a total token count of 802,510. The selected journals include *Applied Linguistics* (SNIP: 2.661), *English for Specific Purposes* (SNIP: 2.249), *International Journal of Corpus Linguistics* (SNIP: 1.211), *Journal of Pragmatics* (SNIP: 1.666), *Journal of Second Language Writing* (SNIP: 2.838), and *TESOL Quarterly* (SNIP: 2.359). The analysis considers complete texts, omitting front matter such as the cover page, Chinese abstract, tables/figures list, table of contents, and appendices.

This study adopts a multi-method analytical framework to investigate metadiscourse practices in academic writing:

(1) Corpus approach

A hybrid corpus-based (deductive) and corpus-driven (inductive) methodology was applied to two comparable corpora: CLMA\_C (50 Chinese EFL learners' MA theses, 804,935 tokens) and ILJA\_C (100 international journal articles, 802,490 tokens). The design prioritized disciplinary rigor by matching publication years (2010–2013) and SNIP metrics (1.211–2.838) for ILJA\_C, while CLMA\_C included theses from 36 Chinese universities. Non-essential sections (e.g., appendices) were excluded to focus on core academic text (e.g., introductions, discussions). The corpus-based approach validated hypotheses using Hyland's (2005) interactional metadiscourse framework, while the corpus-driven method identified emergent patterns (e.g., rhetorical strategies) without predefined categories.

**The principles of selecting keywords for searching in the corpus.**

Typically, corpus-based approaches rely on frequency statistics (such as log-likelihood tests) to identify and select keywords. This statistical method is particularly useful in determining significant words or phrases in ILJA\_C and CLMA\_C, aiding in the investigation of certainty stance adverbs and hedging verbs (e.g., see Table 2; Table 7; Table 10; Table 12).

In contrast, corpus-driven methods adopt a more flexible selection criterion for keywords. Instead of relying solely on frequency statistics, this approach often selects representative words or phrases based on previous related research. The rationale is that a corpus-driven approach allows for a multi-perspective examination of the target linguistic item. Analyzing every possible word would be both time-consuming and impractical, as many may lack relevance or generalizability.

Therefore, in the analysis of maximizers and self-mentions, the study focuses on two representative items: ‘fully’ as an exemplar of maximizers and ‘we’ as a prototypical self-mention.

#### (2) Comparative analysis

Quantitative methods – normalized frequency analysis (per 1,000,000 words) and log-likelihood tests – were employed to statistically compare metadiscourse features (hedges, boosters, attitude markers, etc.) between the corpora. This revealed significant differences in learners’ and experts’ use of certainty markers (e.g., ‘clearly’) and engagement strategies (e.g., ‘we believe’).

#### (3) Academic discourse analysis

Genre analysis contrasted rhetorical norms in theses versus journal articles, while qualitative examination explored how metadiscourse constructs authorial identity. For instance, learners’ preference for procedural explanations (‘we can conclude that’) versus experts’ use of evaluative stance markers (‘we would argue that’) highlighted divergent communicative priorities.

#### (4) CDA

CDA interrogated ideological underpinnings of language choices. For example, experts’ frequent certainty adverbs (‘clearly’) were interpreted as assertions of epistemic authority, whereas learners’ restrained use reflected cultural norms of academic modesty, exposing power dynamics in knowledge production.

#### (5) Thematic analysis

Concordance lines were thematically coded to identify recurrent patterns, such as collaborative knowledge construction (‘you can see’). Cluster analysis revealed how engagement markers varied lexically between corpora while serving similar dialogic functions.

By synthesizing corpus linguistics, comparative statistics, genre analysis, CDA, and thematic coding, this study achieves a triangulated understanding of metadiscourse. The corpus approach provided empirical grounding; comparative

analysis quantified cross-group differences; discourse analyses (academic and critical) contextualized findings within genre and ideology; and thematic coding uncovered latent rhetorical strategies. This multi-layered design ensures replicable insights into how metadiscourse mediates expertise, cultural norms, and disciplinary acculturation in academic writing.

This chapter incorporates these research methods into three principal segments, designed to validate or disprove the initial hypotheses of the investigation. The first part scrutinizes rhetorical hype in Chinese EFL learners' and experts' academic writing through the examination of certainty stance adverbs and maximizers. The second part delves into hedging strategies in Chinese EFL learners' and experts' academic writing, focusing on the investigation of lexical verbs. The third part explores the construction of authorial identity in Chinese EFL learners' and experts' academic writing, analyzing self-mentions.

The approach used to examine metadiscourse in academic writing for this study is based on K. Hyland's (2005) definition of metadiscourse from interactional perspective and the operational definition of it is modified in relation to the research subjects that «the ways in which writers conduct interaction by intruding and commenting on their message, and involve the reader in the text by expressing the writer's positive attitude, emphasizing certainty, withholding commitment and opening dialogue and explicit reference to author(s) to achieve rhetorical hype strategy, hedging strategy and authorial identity construction» [Hyland 2005:49]. And it should be noted that this study utilized corpus analysis tools to extract and identify engagement markers, the fifth subcategory of K. Hyland's interactional metadiscourse framework, from two corpora. The results indicate relatively infrequent usage (905 occurrences in 802,510 words). Additionally, most instances of “we” function as self-mentions rather than engagement markers. Furthermore, second-person pronouns (“you/your/yours”) primarily appear in quoted excerpts rather than as meaningful discourse features,

rendering this subcategory beyond the scope of the present investigation. Therefore, the interactional metadiscourse framework is modified and to ensure the representativeness of the selected data and its relevance to the three rhetorical strategies (rhetorical hype strategy, hedging strategy, and authorial identity construction) examined in this study, a further refinement was conducted within each subcategory combined with other theories, leading to the following classifications:

- **Certainty stance adverbs** belong to the epistemic category, expressing our affirmation of the propositional message and the validity of the propositional content of the text. They not only have semantic meaning but also interpersonal functions, conveying personal commitment and certainty about truth and value judgments [Biber, Johansson, Leech, Finegan 1999; Hyland 2005].

- **Maximizers**, which serve as boosters in academic writing, also play the role of rhetorical hype. They may be perceived as interactional metadiscourse, aiming at engaging the reader within academic discourse and assisting authors in minimizing reader confusion, avoid controversy over propositions, and actively and accurately express and promote their own attitudes [Biber, Johansson, Leech, Finegan 1999; Hyland and Jiang 2021:191].

- **Hedges** are «devices such as ‘possible’, ‘might’ and ‘perhaps’, which indicate the writer’s decision to recognize alternative voices and viewpoints and so withhold complete commitment to a proposition» [Hyland 2005:52].

- **Self-mentions** refer to «the degree of explicit author presence in the text measured by the frequency of first-person pronouns and possessive adjectives» [Hyland 2005:53].

K. Hyland’s classification of interactional metadiscourse served as the foundation for developing a methodology to analyse metadiscourse markers in academic texts produced by Chinese EFL learners and experts. His classification facilitated the collection of contextual examples of metadiscourse, incorporating certain rhetorical strategies among the corpora of this study. Additionally, K.



Hyland's metadiscourse classification was designed for instances that are independent of specific context or discourse. This study introduces the potential application of Hyland's classification within academic discourse, providing a comparative analysis between Chinese EFL learners and experts for the first time.

### **3.2 Rhetorical hype in academic writing: Investigation of certainty stance adverbs and maximizers**

The academic paper, crucial for disseminating scientific knowledge, derives its credibility from the scientific rationality of research findings and their disciplinary value. Traditionally regarded as objective, academic papers eschew the use of rhetorical language to mitigate potential issues associated with over-interpretation [Ochodo 2013; Macleod 2014], as rhetorical language may engender a dichotomy with scientific discourse [Liu 2021]. However, scientific knowledge encompasses not only objective facts but also persuasive elements that shape universal truths through scientists' rhetorical strategies [Gilbert and Mulkay 1984].

In academic circles, researchers utilize a range of rhetorical strategies to highlight the importance and worth of their work. N. Millar (2019) identified a phenomenon wherein authors utilize exaggerated or laudatory language to embellish or promote their academic work, terming it 'rhetorical hype', also known as the 'hying strategy'. This rhetorical approach often employs metadiscourse, such as positive attitude markers and boosters, to achieve its aims. The following section will demonstrate how Chinese English learners and expert authors utilize specific subcategories of positive attitude markers and boosters, namely certainty stance adverbs and maximizers, respectively, to achieve rhetorical effects of exaggerated promotion and argumentation.

### 3.2.1 Investigation of certainty stance adverbs

The concept ‘stance’ was defined by D. Biber and E. Finegan who explained it as an «overt expression of an author’s or speaker’s attitudes, feelings, judgments, or commitment concerning the message» [Biber and E. Finegan 1988: 1]. D. Biber (2006) asserts that stance, as a linguistic expression, serves to communicate the speaker’s or writer’s level of certainty regarding the accuracy of information, as well as their perspective on it. Stance expressions play a significant role in generating ‘rhetorical hype’, as they convey the writer’s viewpoints of certainty and encourage readers to more readily accept the propositional message conveyed in the text [Hyland and Jiang 2021]. Provided herein is an illustrative example ‘The insights gained in this study clearly verify the benefits of the method’, the word ‘clearly’ demonstrates the writer’s certain attitude and amplifies the verification of the benefits of the method. In this context, the certainty stance adverb ‘clearly’ exemplifies the function of rhetorical hype.

Lexical and grammatical patterns for expressing stance are examined using a corpus-based approach, with a focus on academic written registers [Peacock 2015]. Stance is manifested linguistically through the use of verbs [Thompson and Yiyun 1991; Hunston 1994], adverbials [Biber and Finegan 1988; Biber, Johansson, Leech, Finegan 1999], complement clause constructions [Biber, Johansson, Leech, Finegan 1999; Biber 2015], and metadiscourse features [Hyland 2004].

Stance adverbs are one of the main manifestations of stance expressions and can be divided into three categories: epistemic stance adverbs, attitude stance adverbs, and style stance adverbs [Biber, Johansson, Leech, Finegan 1999]. Certainty stance adverbs belong to the epistemic category, expressing our affirmation of the propositional message and the validity of the propositional content of the text. They not only have semantic meaning but also interpersonal

functions, conveying personal commitment and certainty about truth and value judgments [Hyland 2005]. The use of certainty stance adverbs in academic writing indicates the writer's clear and definite attitude, prompts the reader to understand and accept the writer's viewpoint, facilitates the reader's accurate judgment of the propositional message of the academic discourse, and helps construct a harmonious relationship between the writer and the reader so that the reader is clearly aware of the certainty of the information he or she receives. The appropriate use of certainty stance adverbs in academic writing to express stance not only makes explicit the expressed viewpoint but also takes into account the reader's emotion, thus helping the reader to better understand the attitude conveyed in the discourse and thus resonate with the writer.

D. Biber (2006) points out the semantic function of certainty stance adverbs in two aspects: true propositional information and personal attitude. In other words, the certainty stance adverbs in academic writing not only express the ideas clearly, and more importantly, reflects the apparent attitude of the writer in expressing his or her opinion. K. Hyland (2005) also argues that sometimes one proposition has the dual (propositional and commentary) function. This suggests that when reading academic papers we must be aware that the content stated in academic writing not only conveys propositional information but also has evaluative meaning, as a result of the interactive process between the writer and the reader.

Most studies on certainty stance adverbs are based on genre analysis, exploring its usage characteristics from an interpersonal perspective [Hyland 1998; Lancaster 2016]. The main research findings of them are concluded as follows: The distribution characteristics of certainty stance adverbs vary in different registers, genres, and disciplinary texts. They can express the author's judgments of proposition certainty or possibility, reflecting the function of intensifying the tone. In academic discourse, they have interpersonal significance, expressing the author's attitude towards the proposition while also considering

the reader's feelings and engaging in interactive negotiations on the topic. EFL learners face difficulties in stance expression in academic writing, especially in expressing certainty.

Studies on Chinese learners' certainty stance adverbs in academic discourse indicate that learners use certainty stance adverbs less frequently and have insufficient vocabulary richness. They tend to directly express their stance towards the proposition in an authoritative tone [Xu 2015]. These studies on certainty stance adverbs in China are mainly based on Chinese EFL learners' argumentative corpus and native speakers' corpus. The method of interlanguage comparative analysis is used to analyze certainty stance adverbs, and the results show similarities and differences in the characteristics of certainty stance adverbs' usage between learners and native speakers. X. Zhao (2009) pointed out that the main differences are that learners' use of certainty stance adverbs is concentrated on a few items, lacking diversity. They tend to appear in sentence-initial position and lack flexibility in syntactic position. There are relatively few studies on the use of certainty stance adverbs in academic English writing by Chinese FFL learners. F. Pan (2012), based on a corpus of Chinese and international mechanical engineering journal articles, found that Chinese authors used certainty stance adverbs less frequently overall and tended to be colloquial in their choice of vocabulary and sentence structure, and have insufficient understanding of the interpersonal significance of academic discourse, which affects the recognition and acceptance of their papers in the international academic community. These studies indicate that appropriately using certainty stance adverbs to express the author's stance is challenging for Chinese EFL learners due to the diversity of form and semantics and the complexity of syntax. Proper stance expression, especially certainty stance adverbs, in academic writing can construct the evaluative and interactive nature of discourse, indicate the author's attitude and position towards the proposition, involve readers in discourse, guide readers to make judgments and think critically, and help achieve

the purpose of academic communication. Comparing the usage patterns of certainty stance adverbs between academic English writing by Chinese learners and English papers published in authoritative international journals within the same field can elucidate both the hyperbolic skills in stance expression that learners have acquired and the challenges they encounter. This study examines the usage characteristics and difficulties of certainty stance adverbs in Chinese learners' academic writing by comparing their corpora with the corpus of English papers published in authoritative international journals.

This research aims to address the following questions: What specific syntactic and pragmatic rhetorical type characteristics do Chinese MA EFL learners and experts exhibit in their use of high-frequency certainty stance adverbs in academic writing? What factors contribute to the distinctions observed between them?

To answer these research questions, our research will be designed in the following ways. Firstly, the data for this study consists of two electronic corpora of written texts, namely, CLMA\_C and ILJA\_C, as discussed in Section 3.1.

Secondly, to investigate certainty stance adverbs and their hyperbolic function, this study conducted quantitative (corpus-based analysis) and qualitative (interlanguage and discourse analysis) analytical methodology.

Quantitative analysis in this study together with manual contextual analysis focused on all instances of certainty stance adverbs in both corpora. In order to calculate and analyze these instances, WordSmith 5.0 will be used. WordSmith 5.0 is a software program designed for text analysis in corpus linguistics [Scott 2010]. It provides tools for analyzing language patterns, concordance, and collocation in a given body of text in corpus. WordSmith 5.0 is used for two different tasks in this study. The first task is to generate concordance lines involving certainty stance adverbs in both corpora. The frequency distribution of certainty stance adverbs can be obtained in this process. The second is to sort concordance lines to determine their hyperbolic functions in the discourses. Then,

comparisons are made, in terms of structure and function, in order to find out the features of the learners' usage of certainty stance adverbs.

Log-likelihood and Chi-square Calculator 1.0 is a software tool developed by M. Liang (2010). It is designed to calculate log-likelihood and chi-square statistics for linguistic analysis, particularly in corpus linguistics and language studies. In this study, Log-likelihood and Chi-square Calculator 1.0 is used to make comparisons between the occurrences of certainty stance adverbs used by Chinese EFL learners and those by international experts. In the following section, a statistically significant p-value is set at 0.05. If the p-value exceeds 0.05, it indicates no statistical difference between the two groups of research data. However, if the p-value falls within the range of 0.05 to 0.01, we can conclude that there is a statistical difference between the two groups of research data. If the p-value falls within the range of 0.01 to 0.001, it is considered a significant statistical difference. Finally, if the p-value is below 0.001, it is considered an extremely significant statistical difference.

A qualitative analysis was performed in order to make explicit the results of the quantitative analysis. The research subject refers to the category of K. Hyland and J. Milton (1997) and T. McEnery and N. Kifle (2002) and identifies a total of 10 certainty stance adverbs based on semantic classification of the lexical items listed in the academic discourse: "actually, clearly, obviously, really, indeed, certainly, definitely, absolutely, fundamentally, admittedly". The integration of qualitative and quantitative methods can facilitate more explanatory findings. The quantitative analysis identified the frequency of occurrence of certainty stance adverbs in both corpora. The results of the analysis of the frequency of occurrence of certainty stance adverbs in both corpora were used as the basis for interlanguage and discourse analysis on Chinese EFL learners' rhetorical hype in writer-reader interaction and manifest their interpersonal function.

Following the above research design, firstly the software WordSmith 5.0 is employed to generate concordance lines involving certainty stance adverbs in both corpora, manually removing the usage of adverbs that did not indicate certainty stance and obtained frequency information and normalized frequencies for each certainty stance adverbs with the following results.

**Table 2 Frequency statistics of certainty stance adverbs in ILJA\_C and CLMA\_C**

Certainty stance adverbs	CLMA_C		ILJA_C		Chi-square	Significance (p)
	RF	SF	RF	SF		
actually	234	29.07	160	19.94	13.68	0.000 ***
clearly	156	19.38	200	24.92	-5.57	0.018 *
obviously	87	10.81	41	5.11	16.39	0.000 ***
really	64	7.95	93	11.59	-5.45	0.020 *
indeed	40	4.97	149	18.57	-63.20	0.000 ***
certainly	19	2.36	42	5.23	-8.74	0.003 **
definitely	18	2.24	11	1.37	1.67	0.196
absolutely	11	1.37	2	0.25	6.20	0.013 *
fundamentally	4	0.50	13	1.62	-4.79	0.029 *
admittedly	2	0.25	3	0.37	-0.20	0.652
TOTAL	635	78.89	714	88.97	-4.87	0.027 *

Note: 'RF' stands for 'raw frequency'; 'SF' stands for 'standardized frequency (per 10,000)'; The symbol '-' represents the underuse of certainty stance adverbs by Chinese English learners compared to international journal experts; '\*' indicates statistical difference ( $0.01 < p < 0.05$ ); '\*\*' indicates significant statistical difference ( $0.001 < p < 0.01$ ); '\*\*\*' indicates extremely significant statistical difference ( $p < 0.001$ ).

As shown in Table 2, although the use of certainty stance adverbs in CLMA\_C is diverse, the lexical items of them are only dependent on several

choices. In terms of overall usage, the standardized frequency of using certainty stance adverbs in Chinese EFL learner theses is lower than in international journal articles, and there was a statistical difference between the two corpora ( $p=0.027<0.05$ ). Concerning word frequency, Chinese EFL learners and experts differed in the order of high frequency. This suggests that learners basically realize the category of the certainty stance adverbs, while some of them are not yet sufficiently mastered. This finding verifies the studies of K. Hyland & J. Milton (1997). Eight of the ten certainty stance adverbs have difference, among which ‘clearly’, ‘really’, ‘absolutely’ and ‘fundamentally’ have statistical differences ( $0.01<p<0.05$ ); ‘certainly’ has significant differences ( $0.001<p<0.01$ ); ‘actually’, ‘obviously’ and ‘indeed’ have extremely significant differences ( $p<0.001$ ). The top five certainty stance adverbs appearing in high frequency in CLMA\_C were: ‘actually’, ‘clearly’, ‘obviously’, ‘really’ and ‘indeed’, while the top five adverbs appearing in high frequency in ILJA\_C were: ‘clearly’, ‘actually’, ‘indeed’, ‘really’ and ‘certainly’. The consistency of four out of the five frequently occurring certainty stance adverbs indicates that learners are able to use them in their writing to express their confidence in propositional content. The Chi-square test shows that the frequencies of the five certainty stance adverbs appearing at high frequencies in learners’ theses differ from those in international journal articles ( $p<0.05$ ), learners rely more on ‘actually’ to express deterministic stances and use it more frequently than experts ( $p<0.01$ ), but underuse ‘clearly’, ‘really’ and ‘indeed’. Below, we will analyze the usage characteristics of high-frequency certainty stance adverbs.

After analyzing the overall frequency characteristics of certainty stance adverbs, the following section scrutinizes the utilization of three high-frequency certainty stance adverbs when experts and Chinese EFL learners applying hyping strategy, namely ‘actually’ and ‘clearly’, across two corpora.

The certainty stance adverb ‘actually’ was found to frequently collocate with mental verbs by both Chinese EFL learners and experts, with a frequency of



approximately 39% and 73%, respectively (see figure 1). The usage of mental verbs among learners was found to be limited in diversity. Furthermore, the mental verbs followed by experts were predominantly collocated in passive voice (75%) and in active voice (25%), including verbs such as ‘used’ and ‘indicated’. For example,

(1) “...the linguistic resources they had known at a receptive level were **actually** used to achieve a goal in a real-life situation.” (ILJA\_C)

In example (1), expert writer strategically employs the certainty stance adverb ‘actually’ as a **rhetorical hyping strategy** (see 1.4.2). The expert writer **emphasizes the surprising or unexpected nature** of the linguistic resources being employed to achieve a goal in a real-life situation. The passive voice structure (‘were actually used’) adds to **the sense of detachment or objectivity, underscoring the indisputable nature** of the described situation. By collocating ‘actually’ with ‘used’, the expert writer **emphasizes the contrast** between the perceived knowledge of the linguistic resources and their actual application, thereby amplifying the significance of their utilization.

(2) “In some cases, this heterogeneity has a reason because quite different objects of study are **actually** indicated under the same label of ‘keyword’.” (ILJA\_C)

Similarly, in sentence (2), ‘actually’ is used to highlight the **unexpected or surprising revelation** that different objects of study are indicated under the same label of ‘keyword’. The passive voice structure (‘are actually indicated’) lends an air of objectivity to the statement, suggesting that this phenomenon is an undisputed fact. By collocating ‘actually’ with ‘indicated’, the expert writer **underscores the contrast** between the assumed understanding of heterogeneity and the actual indication of different objects of study, thereby **emphasizing** the significance of this observation.

While the **mental verb** collocations utilized by Chinese EFL learners predominantly employed in active voice (83%) and in passive voice (17%), including verbs such as ‘reflect’ and ‘indicate’ (see figure 1).

(3) “According to this metaphor, some linguistic expressions **actually** reflect the distinction between setting and participants.” (CLMA\_C)

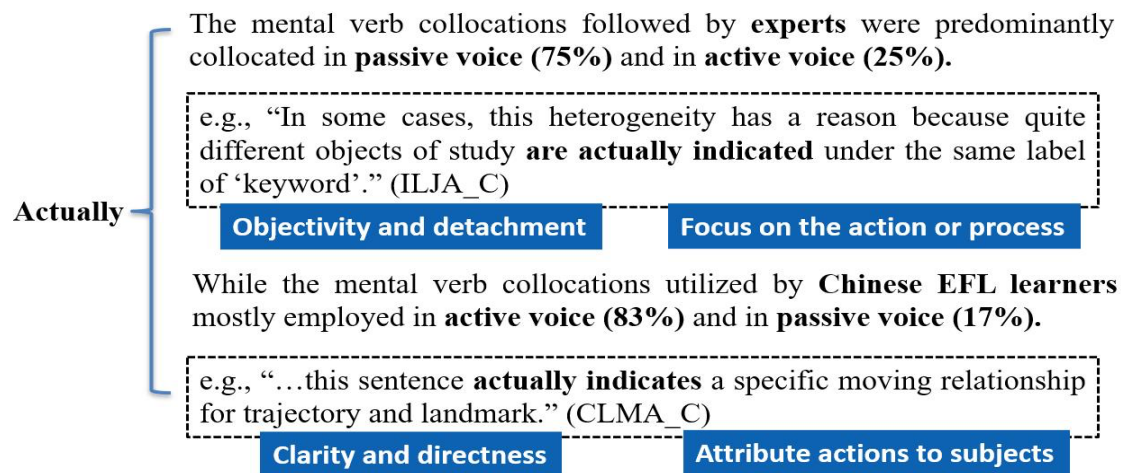
In sentence (3), ‘actually’ is used to **emphasize the unexpected or noteworthy aspect** of linguistic expressions reflecting the distinction between setting and participants. The active voice structure (‘linguistic expressions actually reflect’) lends agency to the linguistic expressions, suggesting that they actively demonstrate this distinction. By collocating ‘actually’ with ‘reflect’, Chinese EFL learners **highlights the contrast** between the assumed understanding of linguistic expressions and their actual ability to reflect the specified distinction.

(4) “...this sentence **actually** indicates a specific moving relationship for trajector and landmark.” (CLMA\_C)

Similarly, in sentence (4), ‘actually’ is employed to **emphasize the surprising or significant nature** of the sentence indicating a specific moving relationship for trajector and landmark. The active voice structure (‘this sentence actually indicates’) suggests that the sentence actively communicates this relationship. By collocating ‘actually’ with ‘indicates’, Chinese EFL learners emphasizes the contrast between the assumed understanding of the sentence and its actual indication of the specified relationship.

The key differences of **rhetorical hyping strategy** between the examples from Chinese EFL learners and experts lie in the syntactic patterns (see figure 1). The possible reasons are Chinese academic writing often emphasizes clarity and directness, which aligns with the use of active voice structures [Xu 2019]. Active voice constructions allow Chinese writers to clearly attribute actions to subjects, making the writing more straightforward and assertive. While in English academic writing, especially in disciplines like linguistics, there is often a

preference for passive voice constructions [Hinkel 2004]. Passive voice structures can create a sense of objectivity and detachment, which is valued in academic discourse. It also allows expert writers to focus on the action or process rather than the agent performing the action.



**Figure 1. The role of ‘actually’ in hyping mental verbs: Syntactic patterns in ILJA\_C and CLMA\_C**

Drawing on the concept of rhetorical hype as defined by Millar et al. (2019), the deployment of the certainty adverb ‘clearly’ is far from a neutral linguistic choice; it is a potent rhetorical strategy that constructs authority, conveys certainty, and negotiates interpersonal positioning in academic discourse. A critical discourse analysis of its collocational patterns reveals a striking contrast between Chinese EFL learners and expert writers. Experts consistently couple ‘clearly’ with concrete referents – such as ‘the figure’ or ‘the table’ – which anchors their claims in explicit, tangible evidence. This syntactic pairing is a classic example of rhetorical hype: by foregrounding the clarity and precision of their results, expert writers transform their empirical data into incontestable facts, thereby amplifying their epistemic authority and aligning their discourse with established academic norms.

In contrast, learners frequently collocate ‘clearly’ with the generic pronoun ‘it’ (in 79% of instances, e.g., “It also shows clearly...”) (see figure 2). This deictic use strips the adverb of its anchoring effect, decontextualizing the

evidentiary basis of their claims. Rather than reinforcing certainty through explicit evidence, this pattern reveals an emerging yet incomplete internalization of academic conventions. Without a concrete referent, the learner's language appears less forceful and persuasive – a gap that not only undermines the argumentative power of their claims but also signals a divergence in ideological positioning.

(10) “It also shows clearly about what kinds of words appear frequently in government documents” (CLMA\_C)

The analysis of specific excerpts further illustrates these differences. In excerpt (10), the learner's use of “clearly” attached only to “it” fails to specify the source or nature of the evidence. The absence of a concrete reference diminishes the impact of the claim and leaves it open to interpretation, thus weakening the perceived validity of the statement.

(11) “The figures clearly show that the differences between Danish and English lectures are not due to a few haphazard instances in our sets of lectures” (ILJA\_C)

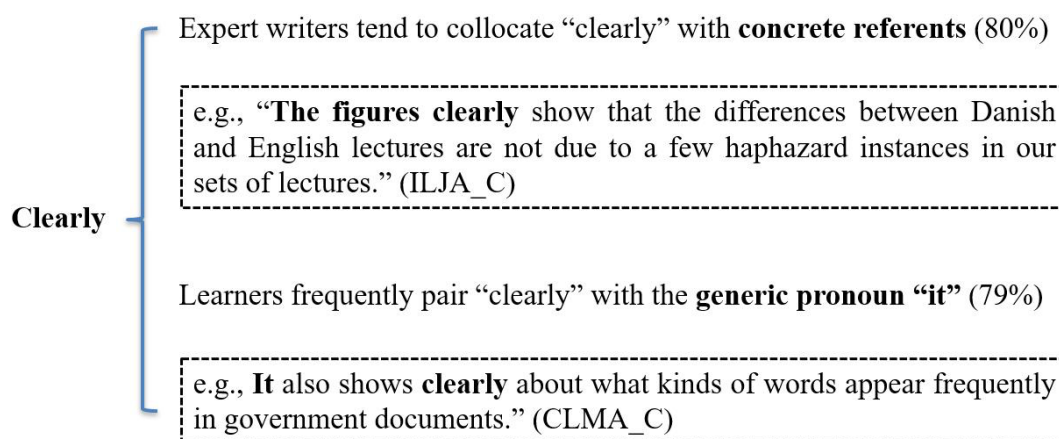
Conversely, in excerpt (11), the expert writer employs “clearly” as a preface to “the figures,” a direct visual concrete referent. This strategic positioning not only highlights the precision of the data but also invokes the norm of evidence-based argumentation, thereby making the claim more persuasive and enhancing the writer's credibility within the academic community.

(12) “A number of tables clearly lay out parallels under such headings as Variability, Variable processes, and Internal constraints on variable units, before going on to address social constraints particular to Deaf communities, and research methods employed in sociolinguistic research on sign languages” (ILJA\_C).

Excerpt (12) further reinforces this pattern. Here, the explicit reference to “tables” immediately ties the adverb to a tangible artifact, ensuring that the claim is seen as systematic, verifiable, and socially validated. This precise collocational

pattern reflects a broader hyperbolic strategy, wherein expert writers use rhetorical hype not merely to embellish but to construct a discourse of undeniable certainty and authoritative knowledge.

Thus, through a CDA lens, the differential deployment of “clearly” is much more than a simple frequency difference – it is an ideological marker that indexes power, identity, and the mastery of academic discourse. Expert writers’ deliberate pairing of “clearly” with specific data not only legitimizes their findings but also embodies the transformative force of rhetorical hype, whereas the learner’s generic usage highlights a gap in their alignment with dominant academic conventions.



**Figure 2. The role of ‘clearly’ in hyping strategy: Collocational patterns in ILJA\_C and CLMA\_C**

To sum up, this study examines the frequency characteristics of all certainty stance adverbs and deeply explores the use of high-frequency of certainty stance adverbs ‘actually’ and ‘clearly’ when writers employing *hying strategy* in Chinese EFL learners’ and experts’ academic writing. Results show that although the diversity of certainty stance adverbs in Chinese EFL learners’ academic writing is varied, the utilization ratio of each lexical item by them is not uniform. Overall, Chinese EFL learners exhibit a lower standardized frequency of using certainty stance adverbs compared to IJAs, with statistical differences observed between the two corpora. Notably, differences in word frequency suggest that while learners grasp the category of certainty stance adverbs, mastery

varies. The study finds significant differences in the usage of certain adverbs, such as ‘clearly’, ‘really’, ‘absolutely’ and ‘fundamentally’, between learners and experts.

Analyzing high-frequency adverbs such as ‘actually’ and ‘clearly’ elucidates that experts collocate ‘actually’ with mental verbs (73% frequency) predominantly in passive voice (75%), enhancing objectivity (e.g., “were actually used”). Learners use active voice (83%) with limited verb diversity (e.g., “actually reflect”), prioritizing directness. Syntactic disparities reflect cultural-academic norms: English favors passive constructions for detachment, while Chinese writing emphasizes active clarity. These patterns underscore how rhetorical strategies are shaped by linguistic conventions, influencing persuasive impact in academic discourse. The certainty adverb ‘clearly’ serves as a rhetorical hype strategy, constructing authority and certainty in academic discourse. Expert writers anchor ‘clearly’ to concrete referents (e.g., “the figures clearly show”), enhancing persuasiveness by linking claims to tangible evidence. In contrast, Chinese EFL learners often pair ‘clearly’ with generic pronouns (e.g., “It also shows clearly”), weakening argumentative impact due to lack of specificity. This divergence reflects differing mastery of academic norms: experts leverage rhetorical hype for epistemic authority, while learners’ usage signals incomplete internalization of disciplinary conventions.

### **3.2.2 Investigation of maximizer ‘fully’**

Maximizers, which serve as boosters in academic writing, also play the role of «rhetorical hype» [Hyland and Jiang 2021:191]. They may be perceived as interactional metadiscourse, aiming at engaging the reader within academic discourse and assisting authors in minimizing reader confusion, avoid controversy over propositions, and actively and accurately express and promote their own attitudes [Hyland 2005]. For instance, ‘As a consequence, findings on

students' beliefs entirely confirm the second hypothesis.', the maximizer item 'entirely' demonstrates the hyperbolic strategy of exaggeration or intensification. It emphasizes the complete and unqualified confirmation of the second hypothesis by the findings on students' belief.

Correct use of maximizers can accurately indicate the speaker's attitude and help persuade others to accept the author's views. In academic writing, appropriate use of maximizers helps authors express the certainty of propositions, enhance the authority of propositions, strengthen the persuasive power of academic discourse, promote academic communication, and achieve the purpose of academic interaction. However, because maximizers express the highest degree of meaning, these words are usually treated as synonyms in the process of learning. Difficulty in distinguishing the commonalities and distinctions among them during the learning process makes it challenging for EFL learners, leading to the inaccurate hyperbolic use of maximizers in academic writing and subsequently impacting the quality of learners' academic writing to some extent [Granger 1998; Lorenz 2014; Zhang 2010].

Therefore, it is essential to analyze the characteristics of maximizers in international journal articles and English academic papers written by EFL learners and compare their differences. Prior research on maximizers has mainly focused on their lexical collocations and grammatical collocations [Granger 1998; Lorenz 2014; Wei 2011; Wang 2007], and pragmatic collocations [Alrajhi, 2019; Özbay and Aydemir 2017]. However, fewer scholars have systematically examined these types of collocations from a macro perspective. This study will use J. Sinclair's Extended Units of Meaning (EUM) model to conduct a multidimensional investigation of maximizers [Sinclair 2004].

Therefore, this study's examination of maximizers mainly involves four tasks: first, extracting the strong collocates around maximizers and categorizing them from a grammatical/colligational perspective; second, categorizing these strong collocates from semantic preference; third, categorizing these strong

collocates from semantic prosody; and fourth, conducting a comparative analysis of the collocational features of Chinese EFL learners and international journal experts within the view of the entire EUM.

Based on these four tasks, this study's research methodology mainly involves corpus-driven analysis, discourse analysis and contrastive interlanguage analysis (CIA). The construction of the EUM model requires bottom-up corpus data driving, and the collocational features of maximizer usage between Chinese EFL learners and international journal experts need to be compared. Explaining the differences between the two necessitates a blend of discourse analysis and CIA.

The research object is to conduct a comparative analysis of maximizers hyping between Chinese linguistic EFL learners and linguistic experts. The research subject aims to investigate the commonalities and distinctions in the EUM of these maximizers within the academic discourse of these two groups.

Maximizers are employed to delineate the utmost intensity within a specified degree range [Altenberg 1991; Quirk, Greenbaum, Leech, Svartvik 1985]. They operate at the semantic level to articulate the absolute degree of a property or state, exemplified by expressions such as 'completely different', 'absolutely endorse', 'totally mix up', etc. On the pragmatic level, as defined by K. Hyland, they function as boosters, constituting an interactional metadiscourse that enables writers to preclude alternative interpretations, forestall conflicting viewpoints, and assert certainty in their assertions [Hyland 2005].

Given that maximizers are often construed as synonyms in traditional dictionaries, the differentiation among them is frequently nebulous, posing a significant challenge for EFL learners in acquisition of the highest degree language.

S. Granger (1998) studied French EFL students' maximizer usage, noting their overuse of 'totally' and 'completely' compared to native speakers. This overuse stemmed from frequent French translational equivalents, driven by a



tendency to make ‘safe bets’. G. Lorenz (2014) explored maximizer collocations in German EFL students’ writing, revealing a tendency to use maximizers in inappropriate contexts. Y. Wei (2011) analyzed maximizer usage in Chinese doctoral dissertations, finding Chinese PhD EFL learners tended to overuse ‘totally’ but underuse ‘entirely’, with congruent collocations indicating a non-native writing style. A. Özbay and T. Aydemir (2017) investigated semantic prosody features of maximizers in Turkish EFL learners’ academic papers, noting incompatibility for some maximizers due to a lack of semantic prosodic awareness in English. M. Alrajhi (2019) explored EFL Saudi students’ use of four maximizers, finding favorable prosody in all, with semantic preferences reflecting emotions and states of mind. T. Peredrienko and E. Balandina (2022) conducted a study on the intensifier ‘очень’ in Russian academic discourse, highlighting its dual role as both a maximizer and a minimizer. They found that ‘очень’ primarily functions as a maximizer, amplifying the significance of lexical units it modifies, with a usage frequency of approximately 91% in academic discourse. The authors also noted a group of synonyms for ‘очень’ in Russian, such as ‘весьма’, ‘крайне’, and ‘исключительно’, commonly employed as maximizers in academic contexts to intensify descriptions and evaluations.

Extended Units of Meaning (EUM) is a corpus-driven approach to language analysis introduced by J. Sinclair (2004). It comprises five components: core, collocation, colligation, semantic preference, and semantic prosody. This methodology integrates form and meaning, providing a multifaceted description of a specific node word. EUM enables the simultaneous examination of lexical, syntactic, semantic, and pragmatic layers, breaking away from previous studies that focused on individual words or fixed phrases as the basic units of analysis. It broadens the scope of semantic research to encompass all relevant forms surrounding the node word. This aligns with J. Sinclair’s advocated maximal approach, which extends the exploration of meaning units beyond single words or fixed phrases.

According to this approach, every unique meaning in language can be associated with its unique lexical form. The generation of lexical meaning comes from two levels: the general meaning of the lexical item and the adjusted meaning generated by co-selection with other lexical items within the meaning unit [Sinclair 2004]. Meaning is not simply divided into lexical meaning and grammatical meaning but is created by continuous text as a whole. Meaning itself is unstable, usually temporary, and negotiable.

Regarding the study of the maximizers in the past, it was rarely studied as a whole EUM unit like scholars such as J. Sinclair did from the perspectives of word frequency, collocation, colligation, semantic preference and semantic prosody. Collocation reflects the collocational words that occur simultaneously with the core (usually a single word or phrase); colligation reflects the grammatical categories that coincide with the core; semantic preference reflects the semantic field of words that occur in conjunction with the core; semantic prosody reflects the semantic function realized by the entire EUM. Studying words from the perspectives of form, meaning, and function is «a useful tool for synonym discrimination and cross-language equivalence research» [Yu and Li 2017:37]. Compared with studies that only focus on word frequency, collocation, and semantic prosody as a single research point, this research model can more comprehensively, objectively, and accurately describe language.

Moreover, the synonymous nature of maximizers presents considerable difficulties for Chinese EFL learners in distinguishing and acquiring them. In studies related to English writing by Chinese learners, the maximizer “fully” is frequently examined, though primarily within the context of argumentative writing [Zhang 2010]. Existing research has predominantly focused on frequency-based corpus analysis of “fully” without accounting for its grammatical, semantic, and pragmatic characteristics [Wei 2011]. Even when examined in the framework of academic discourse, particularly in the context of

rhetorical hyping strategies, the methodology remains largely constrained to corpus-based lexical frequency analysis [Hyland and Jiang 2021].

Therefore, this section takes EUM as a whole from the perspectives of word frequency, collocation, colligation, semantic preference and semantic prosody, and based on international linguistic journal articles and Chinese linguistic MA thesis texts to mainly answer the following two questions: (1) What are the hyping collocational characteristics of maximizer ‘fully’ in Chinese linguistic MA theses and international linguistic journal articles? (2) What similarities and differences exist in the hyping collocational features of maximizer ‘fully’? What are the underlying reasons for these similarities and differences?

Concerning to the above two research questions, this section also conducts research with the aid of the self-constructed Chinese linguistic MA theses corpus (CLMA\_C) and international linguistic journal articles corpus (ILJA\_C). The research process involves the following steps:

Firstly, identifying the collocates of ‘fully’ and classifying its colligation patterns in CLMA\_C and ILJA\_C, respectively.

Secondly, summarizing the semantic preference and semantic prosody of ‘fully’ hyping in ILJA\_C and CLMA\_C, respectively, based on first step.

Thirdly, conducting a comparative analysis on lexical collocation, colligation, semantic preference, and semantic prosody features of the maximizer ‘fully’ in both corpora.

Considering the consistent analytical approaches taken for each maximizer's EUM, this section uses ‘fully’ as a case in point and conducts a thorough investigation of the maximizer ‘fully’ in the following analysis.

To begin with, the collocates of ‘fully’ need to be identified, and then the colligation patterns of ‘fully’ in CLMA\_C and ILJA\_C respectively are classified.

**Table 3. The top 10 significant collocates of maximizer ‘fully’**

<i>CLMA_C</i>					<i>ILJA_C</i>				
Collocates	F(c)	F(n, c)	MI3	LLD	Collocates	F(c)	F(n, c)	MI3	LLD
be	29637	37	11.06	152	be	25357	45	11.84	200
not	2511	14	10.42	94	to	20200	37	11.32	162
understand	696	8	9.85	66	not	3479	19	10.97	120
can	3850	10	8.34	54	that	9806	21	9.91	94
explain	384	5	8.67	44	language	3324	13	9.4	72
they	2363	7	7.51	38	this	5649	12	8.29	52
reflect	268	4	8.22	34	understand	750	7	8.87	50
English	3183	6	6.41	28	develop	491	6	8.81	48
teacher	786	4	6.67	28	more	2519	7	7.12	34
use	4190	7	6.68	28	yet	193	4	8.4	34

Note: F (c) represents the raw frequency of collocates. F (n, c) is the co-occurrence frequency of the node and collocates ( $\pm 5$ ). The co-occurrence frequency, LLD value and MI3 are set to 2, 3.84, 3, the majority of verbs (including be-verb) covers all their inflected form.

Based on Table 3, significant collocates of the maximizer ‘fully’, in *ILJA\_C*, the collocates to the left side of ‘fully’ can be grammatically categorized into two types: infinitive phrases and Be-verbs. To the right side of ‘fully’, most of the collocates are lexical verbs, such as ‘understand’ and ‘develop’, or adjectives, such as ‘acceptable’. Overall, two distinct types of colligation patterns are categorized: ‘infinitive phrases + fully + mental verbs’ and ‘Be-verbs + fully + passive voice of communication verbs/action verbs/happening verbs/evaluative adjectives’.

In *CLMA\_C*, ‘fully’ is frequently collocated with modal verbs and Be-verbs. The prevalence of using modal verbs is attributed to native language transfer, as Chinese learners often utilize ‘能[Néng]/能够[Nénggòu]’ (having an

ability to do something) in Chinese, conveying the author's subjective confidence in interpreting linguistic phenomena. The English equivalents are typically 'can/could/will'. On the right side, collocates resemble those used by experts, involving lexical verbs like 'understand', 'explain' and 'reflect', or adjectives like 'consistent' and 'grammatical'. These form two colligation patterns: 'modal verbs + fully + communication/mental/existence verbs' and 'Be-verbs + fully + passive voice of communication/action/happening/mental/existence verbs/evaluative/relational adjectives'. The first pattern is less common in ILJA\_C, potentially indicating the developmental aspect of Chinese EFL learners and their seemingly congruent colligational patterns.

Subsequently, the collocates of 'fully' need to be identified, and then the semantic preference and semantic prosody of 'fully' hyping in ILJA\_C and CLMA\_C respectively can be summarized.

In ILJA\_C, when 'fully' is paired with expressions like 'require to', 'need to' and 'enable to', indicating obligation and ability, the subsequent collocates on the right side often include mental verbs such as 'understand' and 'develop'. These mental verbs are frequently accompanied by noun phrases like 'the complexity of' and 'factual information'.

Experts strategically employ the maximizer 'fully' in their research papers to accentuate or hype a high degree of involvement. The semantic preference of 'fully' is evident in its collocation with expressions indicating obligation and ability on the left side, such as 'require to', 'need to' and 'enable to', as well as with mental verbs like 'understand' and 'develop' on the right side. The use of accompanying noun phrases, like 'the complexity of' and 'factual information', highlights the author's preference for emphasizing the thoroughness and completeness of mental processes, aligning with the maximizer 'fully'. This semantic preference conveys the idea that when grappling with intricate theories or data, a comprehensive and thorough approach is essential. It underscores the

notion that the actions described by the mental verbs should be executed to their fullest extent.

The semantic prosody of ‘fully’ in this context is positive, contributing to a tone that is commendatory. The author’s choice of words creates a positive association with the maximizer ‘fully’, suggesting that a thorough engagement with complex theories or data is not only recommended but also essential and praiseworthy.

In summary, the semantic preference of ‘fully’ here emphasizes thoroughness and completeness in mental processes, especially in understanding and developing approaches to complex theories or data. The semantic prosody aligns with a positive and commendatory tone, reinforcing the notion that a comprehensive approach is not only advisable but also imperative and commendable.

When ‘fully’ is combined with copular verbs on the left, the predominant right-side collocates consist of passive forms of lexical verbs pertaining to communication, action, occurrence, and evaluative adjectives, such as ‘explained’, ‘supported’, and ‘acceptable’, etc. This pattern suggests a semantic preference, conveying ‘The theory or data is thoroughly explained, substantiated, and appropriately accepted, aligning with the intended state of the author’, indicating a positive stance. Overall, the semantic preference of ‘fully’ in international journal articles can be summarized as ‘facing complex theories or data, the author must fully comprehend and articulate them. This ensures readers readily accept their perspectives, reaching the expected state of the author’, presenting an objective and positive semantic prosody.

Compared to ILJA\_C, CLMA\_C shows a reduced collocation frequency between ‘fully’ and infinitive phrases but an increased frequency with modal verbs. Notably, many of these modal verbs, like ‘can’, ‘could’, ‘may’, ‘will’, are permission or volition verbs. The right-side lexical verbs, often co-occurring with these modal verbs, include communication, mental, and existential verbs like

‘explain’, ‘understand’, ‘reflect’. Concordance lines suggest these verbs often pair with noun phrases related to ‘theory, utterances, evaluative features’, indicating a semantic preference of ‘subjectively desiring to fully interpret the theory and understand the meaning of language’, expressing the author’s positive subjective initiative.

The semantic preference of ‘fully’ in CLMA\_C suggests a subjective desire for a thorough interpretation of theories and understanding language nuances, reflecting the author’s positive initiative. The semantic prosody maintains a positive and affirmative stance. In contrast to the rhetorical hype utilized by learners, experts tend to adopt a commendatory tone, employing rhetorical devices to amplify descriptions for the purpose of praise, admiration, or approval, which is distinct from the affirmative tone in CLMA\_C that focuses on confirmation or agreement.

When ‘fully’ is collocated with a Be-verb to the left, the types of verbs and adjectives to the right resemble those found in international journal articles, demonstrating a semantic preference of something to be fully utilized and reach expectation state. But the subject of the Be-verb is typically a relatively vague concept. For example:

(13) “...and thus the public expectations are fully satisfied.” (CLMA\_C)

(14) “It is claimed that the theoretical statements are fully consistent with...” (CLMA\_C)

In contrast, the subject of international journal experts is usually a more concrete concept. For example:

(15) “Silva and Leki’s (2004) description of L1 composition’s view of language, is most fully articulated by Berlin (1988).” (ILJA\_C)

Chinese EFL learners, using ‘fully’, prefer vague subjects like ‘public expectations’, indicating fulfillment or general statements. This reflects a tendency toward generalized language use. In contrast, experts pair ‘fully’ with

concrete subjects, emphasizing maximization in specific contexts, making its usage more specific.

In Chinese EFL learners' examples, the semantic prosody of 'fully' may convey a positive tone, indicating satisfaction or accomplishment. In contrast, international journal experts emphasize completeness in specific contexts, reflecting a rigorous and in-depth approach.

Chinese EFL learners often opt for ambiguous concepts as subjects when employing 'fully', suggesting a preference for generalizations and hinting at a potential divergence in their grasp of precision and correctness in academic texts.

This may stem from differences in language learning stages and familiarity with specific terms and concepts.

In contrast, experts prioritize presenting the degree of 'fully' in specific contexts with concrete subjects, indicative of a higher requirement for precision and specificity in academic expression.

Table 4 and Table 5 provide specific information of EUM of maximizer 'fully' hyping in ILJA\_C and CLMA\_C respectively.

**Table 4. The EUM of 'fully' in ILJA\_C**

<b>Collocation</b>	require to, needed to, enable to, etc.		understand, develop, grasp etc.
	be, is, are, was		explained, supported, acceptable, etc.
<b>Colligation</b>	infinitive phrases	<b>fully</b>	mental verb
	be-verbs		passive forms of lexical verbs pertaining to communication, action, occurrence, and evaluative adjectives
<b>Semantic preference</b>	the author intends to convey the idea that when engaging with complex theories or data, it is crucial to approach them with a comprehensive and thorough mindset.		
	the theory or data is thoroughly explained, substantiated, and appropriately accepted, aligning with the intended state of the author		
<b>Semantic prosody</b>	with a positive and commendatory tone		
	objective and positive semantic prosody		



**Table 5 The EUM of ‘fully’ in CLMA\_C**

<b>Collocation</b>	can, could, may, will	<b>fully</b>	understand, explain, reflect
	be, is, are, was		utilized, reflected, grammatical, consistent
<b>Colligation</b>	modal verb		communication, mental or existence verbs
	be-verb		passive forms of lexical verbs pertaining to communication, action, occurrence, and evaluative adjectives
<b>Semantic preference</b>	learners’ subjective desire for a thorough interpretation of theories and understanding language nuances		
	something to be fully utilized and reach expectation state and generalizations or a potential variation in their understanding of specificity and accuracy in academic writing		
<b>Semantic prosody</b>	a positive and affirmative stance		
	objective and positive semantic prosody		

In summary, this section mainly examines the EUM of maximizer ‘fully’ hyping in CLMA\_C and ILJA\_C. The study found that overall the EUM of high-frequency ‘fully’ in both corpora have significant differences in their colligation, semantic preference and semantic prosody. At the formal level, compared to international journal experts, Chinese EFL learners tend to use collocates and colligational forms that are rarely used by international journal experts. This phenomenon may be influenced by native language transfer and embodies their interlanguage developmental features. Concerning semantic preference, Chinese EFL learners tend to hype their subjectively desiring to fully interpret the theory and understand the meaning of language. While international journal experts tend to hype their engagements with approaching complex theories or data in a comprehensive and thorough mindset. In the realm of semantic prosody, experts employ a commendatory tone, while Chinese EFL learners manifest a positive and affirmative tone, which does not effectively showcase the use of hyping

strategies. In teaching, teachers should boost learners' awareness of form-meaning connections. Learners must grasp each maximizer's EUM both formally and semantically, enhancing academic writing quality.

### **3.3. Hedging strategy in academic writing: An investigation of lexical verbs**

Hedging is a linguistic phenomenon that serves various essential functions in academic discourse, serving as metadiscourse devices that enable authors to present their claims with different degrees of caution, express their viewpoint on statements made, and engage in a dialogue with the reader [Hyland 1996]. For example, 'The results suggest that excessive use of quotation fragments could be a crucial phase in the evolution of...' hedging verb 'suggest' can assist writers in expressing uncertainty regarding knowledge and displaying reduced commitment to a proposition. This reflects the author's true depth of knowledge and comprehension regarding the outcomes achieved. Additionally, hedges can moderate a claim, presenting it as a viewpoint instead of an absolute fact [Hyland 1998; Salager-Meyer 1997]. This is particularly valuable in situations where categorical statements might attract potentially face-threatening criticism and opposition. Hedging can also empower the author to express humility and respect towards the audience [Salager-Meyer 1997]. It not only helps to avoid unwanted conflict but also demonstrates a willingness to engage in dialogue and exchange of opinions. Thus, the crucial interactional and social roles that hedging can serve, along with its function in expressing subtle distinctions in certainty and commitment, have been well documented [Poos and Simpson 2002]. This explains why a certain amount of hedging has become standard practice across various disciplines, enabling authors to follow the writing style expected by their discourse community.

Hedging can have a notably significant impact on refining language in language-dependent fields like linguistics. Nonetheless, its capacity to add subtle

meanings to a proposition can present challenges in its interpretation for novice or aspiring members of a discourse community, particularly those from different linguistic backgrounds. If handling subtle language is already challenging for individuals within their own linguistic, academic, and general cultural context, it becomes notably more difficult for individuals from entirely different backgrounds, trained in diverse linguistic educational systems, and who are not native English speakers. This situation arises when non-native international students pursue a higher education linguistics degree in English-speaking countries like the UK, the US, etc. or in Master's or doctoral programs conducted in English like the English-taught Master's or PhD programs offered by RUDN University.

Interpreting hedges effectively in writing can be particularly challenging for advanced non-native speakers [Abbuhl 2006]. This issue may arise from several factors, including unfamiliarity with potential hedging items and a limited comprehension of the sociopragmatic and disciplinary conventions governing the use of hedges [Rosa 2012; Tessuto 2011]. Moreover, the lack of specific teaching activities designed to help non-native students identify and understand hedging correctly may further complicate the issue. [Hyland 2003; Wishnoff 2000]. Ultimately, the failure to interpret hedging can lead to second-language speakers' inability to grasp a native speaker's intended meaning [Fraser 2010].

To date, extensive research has been conducted on hedging in academic discourse. From a cross-disciplinary perspective, T. Varttala (2001) analyzed the use of hedging in popularized articles compared to research articles in three fields: economics, medicine, and technology. M. Takimoto (2015) examined their occurrence and function in journal papers within humanities and social sciences. In terms of cross-language comparative analysis, H. Kreutz and A. Harres (1997) carried out an in-depth study on the dispersion and role of hedging in English and German academic writing. Their findings indicated that while hedging in English texts is used to soften arguments and assertions, in German writing, it often serves

to assert authority. I. Vassileva (2001) compared the use of commitment and detachment in English and Bulgarian linguistic research texts, revealing marked variations in the dispersion of hedging and boosting elements across different sections, potentially resulting in misunderstandings in cross-cultural communication. The findings emphasize the importance of understanding cultural-specific features when using English as the international language for academic communication. P. Cabanes (2007) presented an approach to contrastively study hedges in English and Spanish architecture project descriptions. The analysis suggests a diverse range of lexico-grammatical features indicating various hedging strategies. These strategies serve three rhetorical functions: expressing deference and politeness towards the audience, protecting the authors from potential negative consequences, and taking into account the necessary exactitude in texts influenced by cognitive factors and norms within the architectural discourse community. O. Boginskaya (2022) examines the diachronic changes in hedging elements in Russian academic discourse, analyzing 112 research article abstracts published between 2008 and 2021, revealing shifts in hedging strategies, with a transition from diminishing authorial presence to signaling methodological limitations and research result inaccuracies over time. G. Hu and F. Cao (2011) conducted a comparative study on hedges and boosters of abstracts in Chinese and English articles. The distinct use of them jointly influences authorial certainty and confidence, leading to implications for rhetorical strategies, epistemological beliefs, English language proficiency, and supporting evidence in academic texts. While there is a scarcity of research that examines and compares English language learners, the majority of existing studies involve comparisons between native speakers and English learners [Hyland and Milton 1997; Lee and Deakin 2016; Takimoto 2015]. Surprisingly, there are significantly fewer studies [Aull and Lancaster 2014; Haufiku and Kangira 2018] that specifically address novices and proficient linguistic writers. This study will explore and compare the collocational resources

of lexical verbs hedging in Chinese linguistic EFL learners' academic writing with linguistic experts' counterpart and obtain their contextual features of lexical verbs hedging in academic writing.

Hedging, commonly linked to modal verbs, is able to likewise be extended to lexical verbs, as suggested by K. Hyland (1998). Lexical verbs are the most frequently utilized hedging expressions in scientific articles. Despite being acknowledged as crucial rhetorical tools for achieving pragmatic politeness and employing hedging strategies, lexical verbs often receive inadequate attention within academic discussions [Perez-Llantada 2010]. Lexical verbs offer a straightforward method of encoding epistemic subjectivity, thereby functioning to temper commitment and assertiveness. Lexical verbs exhibit rhetorical flexibility in situations where direct assertions might not be the most efficient way to communicate [Hyland 1998]. These verbs can be classified according to how they indicate the writer's level of certainty regarding a proposition's truth: speculative verbs such as 'propose' and 'consider' express subjective viewpoints; deductive verbs like 'conclude' and 'deduce' signify inferences drawn from established facts; quotative verbs such as 'claim' denote information from secondary sources; and sensorial verbs like 'feel' or 'look' reflect observations based on sensory perception.

Regarding the aforementioned lexical verb hedging research category, this part concentrates the representativeness of the corpora, offering an approximate yet informative glimpse into hedging frequency and function within a specific set of specialized texts from China during 2011-2013. The adoption of a synchronic approach was driven by the intention to examine contemporary hedging usage, while avoiding analysis of its historical evolution. The data comprises two electronic corpora: CLMA\_C and ILJA\_C, as discussed in Section 3.1.

To investigate lexical verbs and their hedging rhetorical functions, this study conducted quantitative (corpus-driven analysis) and qualitative (discourse analysis) analyses.

The quantitative analysis in this study, together with manual contextual analysis, focused on all instances of lexical verbs in both corpora. Firstly, AntConc 4.1.4 is applied to draw concordance lines involving lexical verbs in both corpora. Secondly, the lexical verbs hedging in the two corpora were identified manually according to the categories of lexical verbs functioning as hedges [Hyland 1998; Vass 2017]. Thirdly, the difference in word-count between CLMA\_C and ILJA\_C is normalized per 1,000,000 tokens (TPM). Fourthly, Log-likelihood and Chi-square Calculator 1.0 is used to make comparisons between the occurrences of lexical verbs used by Chinese linguistic EFL learners and those by linguistic experts. In this study, a p-value of 0.05 was established as the threshold for statistical significance. [Liang 2010]. If the p-value is less than 0.05, we can conclude that the results are statistically significant. Instead, if the p-value is equal or larger than 0.05, it will indicate that there is no significant difference between two research data. To obtain the collocational information of lexical verbs in hedging, we utilized AntConc's Collocates function, which calculates the collocational strength considering with Log-likelihood value to identify significant collocates for each lexical verb. The calculated results are shown in Table 8, Table 11, and Table 13. By considering these significant collocates, we conducted a contrastive discourse analysis.

Qualitative analysis was performed in order to make explicit the results of the quantitative analysis. The research subject referred to the category of lexical verbs, identifying a total of 40 high-frequency lexical verbs listed in the Appendix [Hyland 1998; Vass 2017]. The integration of qualitative and quantitative methods can facilitate more explanatory findings. The quantitative analysis identified and obtained statistical information of lexical verbs hedging in both corpora. These numerical results were used as the basis for discourse and comparative analysis on lexical verbs' hedging strategy concerning with Chinese linguistic EFL learners and linguistic experts. The study will outline its research questions in the following manner:

(1) What are the discrepancies in frequency mapping of high-frequency lexical verbs used for hedging in academic writing between Chinese linguistic EFL learners and linguistic experts?

(2) What are the collocational features of high-frequency lexical verbs hedging in linguistic experts' academic writing, and what are the heterogeneous collocational features found in Chinese linguistic EFL learners' academic writing?

Firstly, in response to the first research question, the upcoming section will analyze the 40 high-frequency lexical verbs categorized by K. Hyland as hedging markers potentially employed for academic purposes [Hyland 1998; Vass 2017]. AntConc will be utilized to count the frequency of lexical verbs in each category. Additionally, the concordance function of AntConc will be employed to obtain the raw frequencies of lexical verbs in different categories, and their frequencies will be calculated in tokens per million (TPM). Based on these statistics, the software Log-likelihood and Chi-acquire Calculator 1.0 will be used to compare the discrepancies and calculate their Log-likelihood and Significance (p) values. Detailed data information are shown in Table 6.

**Table 6. Frequency distribution of lexical verbs across two corpora**

Verb category	CLMA_C		ILJA_C		Log-likelihood	Significance (p)
	RF	TPM	RF	TPM		
Speculative	1603	2047.59	2879	3686.79	-371.57	0.000
Deductive	481	614.41	490	627.48	-0.11	0.743
Quotative	594	758.75	490	627.48	9.73	0.002
Sensorial	762	973.34	360	461.01	146.27	0.000
TOTAL	3454	4411.97	4242	5432.21	-82.83	0.000

Notes: "RF" stands for "raw frequency", representing the actual count of lexical verbs in each category within the corpora. "TPM" stands for "frequencies in tokens per million", which indicates the frequency of lexical verbs per million words. "-" indicates the underuse of

hedging lexical verbs in the CLMA\_C corpus, suggesting that these verbs are used less frequently compared to the ILJA\_C corpus being analyzed.

Table 6 describes the overall usage characteristics of different types of lexical verbs in two corpora. Speculative, deductive, quotative and sensorial verbs are extensively used in both corpora but exhibit differences in frequency. Specifically, the TPM of lexical verbs used by Chinese linguistic EFL learners is significantly lower than that of linguistic experts ( $P\text{-value} = 0.000 < 0.05$ ), with frequencies of 4411.97 and 5432.21 TPM, respectively. The proportions of them in CLMA\_C are 46.41%, 13.93%, 17.60%, and 22.06%, respectively, while in ILJA\_C they are 67.87%, 11.55%, 12.09%, and 8.49%, respectively. From the distribution of the four types of lexical verbs in the two corpora, speculative verbs exhibit the highest frequency of use in both corpora, deductive verbs are used least frequently in CLMA\_C, and sensorial verbs are used least frequently in ILJA\_C. The statistical analysis indicates that the frequency allocation of deductive verbs does not differ remarkably between the two corpora. ( $P = 0.743 > 0.05$ ), but significant disparity exists in the frequency allocation of speculative, quotative, and sensorial verbs ( $P = 0.000 < 0.05$ ). The above data analysis indicates that both Chinese linguistic EFL learners and linguistic experts can use different types of lexical verbs in academic English writing. However, their differences suggest that there is still a gap in the rhetorical versatility of hedging between Chinese linguistic EFL learners and linguistic experts in academic writing.

Given the substantial variances in the frequency allocation of speculative, quotative, and sensorial verbs between the two corpora, the next two sections primarily examine these differences, as well as the collocational patterns of these verbs with particular lexical items.



### 3.3.1. An investigation of speculative verbs

Speculative verbs are indexed separately using AntConc, and the raw frequency and TPM values are obtained accordingly. Next, Log-likelihood and Chi-square Calculator 1.0 are applied to obtain distinct frequency characteristics for each speculative verb in CLMA\_C and ILJA\_C, with the results detailed below.

**Table 7. Frequency distribution of speculative verbs across two corpora**

Speculative verbs	CLMA_C		ILJA_C		Log-likelihood	Significance( <i>p</i> )
	RF	TPM	RF	TPM		
indicate	492	628.46	611	782.43	-13.17	0.000
consider	325	415.14	612	783.71	-90.07	0.000
suggest	224	286.13	541	692.79	-136.21	0.000
propose	155	197.99	69	88.36	33.67	0.000
think	131	167.33	318	407.22	-80.78	0.000
expect	115	146.90	207	265.08	-26.89	0.000
believe	91	116.24	156	199.77	-17.47	0.000
argue	30	38.32	252	322.71	-200.36	0.000
claim	30	38.32	87	111.41	-29.13	0.000
doubt	5	6.39	5	6.40	-0.00	0.997
contend	3	3.83	11	14.09	-4.88	0.027
speculate	2	2.55	10	12.81	-5.84	0.016
TOTAL	1603	2047.59	2879	3686.79	-371.57	0.000

Notes: “RF” stands for “raw frequency”, representing the actual count of lexical verbs in each category within the corpora. “TPM” stands for “frequencies in tokens per million”, which

indicates the frequency of lexical verbs per million words. “-” indicates the underuse of hedging lexical verbs in the CLMA\_C corpus, suggesting that these verbs are used less frequently compared to the ILJA\_C corpus being analyzed.

As illustrated in Table 7, Chinese linguistic EFL learners have some similarities with linguistic experts. Chinese linguistic EFL learners are overall in line with the diverse types of speculative verbs in experts’ writing instead of relying on a few ones. Specifically, they have similarities in the proportion of some high-frequency speculative verbs, ‘indicate’ [CLMA\_C:30.70% (492 out of 1063); ILJA\_C:21.22% (611 out of 2879)], ‘consider’ [CLMA\_C:20.27% (612 out of 2879); ILJA\_C:21.26% (492 out of 1063)], ‘suggest’ [CLMA\_C:13.97% (224 out of 1063); ILJA\_C:18.79% (541 out of 2879)]. Regarding to the differences, speculative verbs in ILJA\_C are more frequently used than that in CLMA\_C. Besides they have significant differences in frequency ( $p=0.00 < 0.05$ ). The individual differences of using speculative verbs can be concluded as follows: (1) Chinese linguistic EFL learners use ‘indicate’ most frequently (46.28%, 492 out of 1603), while experts use ‘consider’ most frequently (21.26%, 612 out of 2879); (2) Compared with experts, Chinese linguistic EFL learners significantly use ‘argue’ and ‘suggest’ less ( $LL=-200.36, p<0.05$ ;  $LL=-136.21, p<0.05$ ), and use ‘propose’ more ( $LL=33.67, p<0.05$ ). This suggests that while Chinese EFL learners possess the ability to employ a range of speculative verbs similarly to experts, notable distinctions exist in their selection of specific speculative verbs. In specific contexts, how do they use speculative verbs to achieve the communicative purpose of employing hedging? What are the specific collocational features? We take ‘suggest’ as an example, because Table 7 shows that ‘suggest’ is used frequently and has significant differences in both corpora.

**Table 8. Collocational information of ‘suggest’ in CLMA\_C and ILJA\_C**

CLMA_C				ILJA_C			
Collocate	FreqL	FreqR	Likelihood	Collocate	FreqL	FreqR	Likelihood
that	10	182	383.932	that	15	458	882.548

results	15	0	27.671	results	37	1	80.711
further	8	5	23.663	this	90	33	80.613
name	4	1	20.429	may	8	34	44.161
in	22	29	20.412	findings	22	0	43.908
should	0	15	20.362	might	6	16	36.566
findings	7	3	18.161	evidence	16	0	26.999
author	11	0	16.511	observations	6	1	19.379
future	3	3	13.702	in	85	53	18.233
it	34	3	13.334	some	9	14	15.147

Due to space limitations, only the top 10 significant collocates of ‘suggest’ (ranked by Likelihood, refer to Table 8) are presented here. Table 8 displays the significant collocates of ‘suggest’ in CLMA\_C, including ‘that’, ‘results’, ‘further’, ‘name’, ‘in’, ‘should’, ‘findings’, ‘author’, ‘future’ and ‘it’. Similarly, the significant collocates of ‘suggest’ in ILJA\_C are different, consisting of ‘that’, ‘results’, ‘this’, ‘may’, ‘findings’, ‘might’, ‘evidence’, ‘observations’, ‘in’ and ‘some’. Notably, ‘that’ emerges as the strongest significant collocate in both CLMA\_C and ILJA\_C. Additionally, the concordance lines indicate that ‘that’ predominantly appears in R1 (first position on the right) of ‘suggest’ with a co-occurrence probability of 94.79% (182 out of 192) in CLMA\_C and 96.83% (458 out of 473) in ILJA\_C. Additionally, Table 9 depicts the co-occurrences in L1 and L2 (first and second position on the left). The occurrence of lexical bundles with the structure ‘X X suggest\* that’ was calculated.

**Table 9. The co-occurrences in the L1 and L2 with ‘suggest/ suggests/ suggested that’**

CLMA_C	Frequency	ILJA_C	Frequency
It is suggested that	16	the results suggest that	11

he suggests/suggested that	7	the findings suggest that	9
the results suggest that	5	this result suggests that	6
this finding suggests that	3	these results suggest that	5
the author suggests that	3	These findings suggest that	4
as is suggested by	3	it is suggested that	4
further suggests that	3	this study suggest that	3
this study suggests that	2	the data suggest that	3
the network suggests that	2	studies have suggested that	3
sources may suggest that	2	some evidence suggests that	2
TOTAL	46	TOTAL	50

Notes : frequency $\geq$ 2

Table 9 shows that Chinese linguistic EFL learners exhibit comparable lexical diversity to experts. In ILJA\_C, most of the collocates preceding ‘suggest’ are ‘result’ or ‘results’ (22 out of 50, 44.00%) and ‘finding/findings’ (13 out of 50, 26.00%), suggesting that experts commonly employ the pattern ‘results/findings suggest that’ to present research results frequently. In contrast, the co-occurrence probabilities of ‘result/results’ and ‘finding/findings’ preceding ‘suggest’ in CLMA\_C are merely 10.87% (5 out of 46) and 6.52% (3 out of 46), respectively. Instead, the more frequently employed pattern is ‘It is suggested that’ (16 out of 46, 34.78%). For example,

(16) “It is suggested that lexical chunks teaching should be conducted in contexts and the teachers’ mindset should be changed in teaching.” (CLMA\_C)

Impersonal constructions serve as a type of hedging strategy [Hyland 1998; Vass 2017]. Experts frequently employ phrases such as ‘result / finding / study / data / evidence / observation / sources / it + suggest’ to detach themselves from their assertions, allowing them to report findings without explicitly disclosing

their personal stances. Among these, the ambiguity of ‘it’ is relatively low. Although ‘it’ can objectively state viewpoints or facts and reduce the subjectivity of the author’s proposition, its excessive use leads to unclear source information and diminishes the authority and credibility of the reported results, which is not conducive to establishing a strong connection between the author and readers. As a result, this language pattern is less prevalent in ILJA\_C. However, the prevalent use of the ‘It is suggested that’ language pattern by Chinese linguistic EFL learners indicates their awareness of the scientific objectivity in linguistic research paradigms. It also reflects their recognition of their low authority and hesitation to intervene in propositions. To enhance the objectivity of their papers, they attempt to minimize their subjective involvement in proposition discussions.

Additionally, hedging strategies are crucial in restating viewpoints during literature reviews. Chinese linguistic EFL learners commonly use personal pronouns to individually review the cited literature (see Table 9). For example,

(17) “Cheng (2002) defines them more critically. He suggests that course-books should be divided into two categories.” (CLMA\_C)

Experts typically adopt a comprehensive approach by categorizing the cited literature in advance and then employing ‘shell nouns’ (e.g. finding, result) (see Table 9) to concisely summarize the supporting information from previous studies in an abstract manner, connecting it seamlessly to the subsequent text [Schmid 2012]. For example,

(18) “Taken together, these findings suggest that focused, metalinguistic feedback serves to improve learners’ grammatical accuracy.” (ILJA\_C)

In example (18), experts employ the phrase ‘these findings suggest’ to succinctly summarize previous studies on metalinguistic feedback and convey their benefits to the readers. In the aforementioned examples, both personal pronouns and shell nouns are identified as hedging techniques. However, compared to novices’ use of personal pronouns for individual reviews, employing shell nouns to summarize viewpoints can establish a more robust connection

between ideas, reduce readers' cognitive processing time, enhance the credibility of the viewpoints, and improve the acceptance of the presented ideas.

The reason why Chinese linguistic EFL learners tend to express themselves in this manner can be attributed to two factors. Firstly, Chinese academic writing is often influenced by the Confucian value of 'knowledge telling' [Hyland 2005]. As a result, Chinese linguistic EFL learners may rely excessively on retelling information rather than drawing conclusive statements. Secondly, Chinese linguistic EFL learners may possess relatively basic skills in citation and literature review.

Moreover, experts typically employ first-person pronouns in self-referring or speculative language when presenting their own viewpoints [Vass 2017]. For instance, phrases like 'I/we suggest' are utilized to convey the author's subjective opinion. Conversely, Chinese linguistic EFL learners are influenced by Chinese academic norms when using self-referring or speculative language. For example,

(19) "Finally, the author suggests that the effect of the mother tongue on acquisition of the target form should be considered in the process of teaching activities." (CLMA\_C)

In example (19), Chinese linguistic EFL learners aim to convey their belief that the mother tongue impacts the acquisition of the target language form in teaching activities, using the phrase 'the author suggests' instead of 'I suggest' (see Table 9). This choice is influenced by Chinese academic norms, which place a strong emphasis on objectivity in academic writing, requiring authors to minimize the use of subjective language that reveals their personal identity [Lu 2020]. Chinese EFL learners might also be limited by their apprentice status, and consequently, they may seek to mitigate the prominence of their self-identity in order to gain recognition from the defense panel. As a result, authors are advised to avoid first-person expressions and instead use phrases like 'the author believes' or 'this article believes'. Chinese linguistic MA learners, being writing novices,

are deeply influenced by these academic norms when constructing their author identity.

Furthermore, in ILJA\_C, experts frequently utilize probability modal verbs like ‘may’ and ‘might’ in that-clauses (see Table 8). This can be observed in Table 8, where ‘may’ and ‘might’ are significant collocates of ‘suggest’, co-occurring 50 times (with a frequency of 64.03 TPM). These modal verbs primarily convey the author’s viewpoint on the relatively uncertain commitment of the proposition. For example:

(20) “The results suggest that these types of benefits may not happen concurrently for the same individuals; the type of benefits received may depend on an individual’s approach to writing and the factors considered when completing the genre-based tasks.” (ILJA\_C)

(21) “... the findings of this study suggest that the SFL approaches might also be beneficial to less proficient FL learners.” (ILJA\_C)

The use of probability modal verbs ‘may’ and ‘might’ serves two purposes. Firstly, it indicates that experts are cautious when reporting study results, allowing room for negotiation and moderating the accuracy and reliability of the proposition. In example (20), the use of the probability modal verb ‘may’ suggests that there is a possibility that the types of benefits may not happen concurrently for the same individuals. The type of benefits received may depend on individual factors and approaches to writing.

Example (21) employs the modal verb ‘might’ to indicate that while SFL genre pedagogies have historically been regarded as most beneficial for advanced learners, the results of this study propose that they could also prove advantageous for FL learners who are less proficient. The use of ‘might’ indicates that this is not a definitive conclusion but rather a possibility supported by the evidence presented in the study. By using this hedging language, the author can present a more objective and persuasive argument that considers the potential limitations or alternative interpretations of the findings.

However, the evidence in CLMA\_C reveals that Chinese linguistic EFL learners tend to use the obligatory modal verb ‘should’ 15 times (with a frequency of 19.16 TPM) (see Table 8) in ‘suggest that-clause’, whereas ‘may’ and ‘might’ appear only 7 times (with a frequency of 8.94 TPM) in the same context. Conversely, modal verb ‘should’ is rarely used in experts’ academic writing. For example,

(22) “Therefore, the results of previous SLA studies suggested that research in this field should study the two separate processes individually and sequentially.” (CLMA\_C)

(23) “This finding suggests that all subjects should review the target forms on time in case they overuse or forget them.” (CLMA\_C)

The selection of modal verbs in reporting results suggests that experts often utilize low probability modal verbs to present their findings. On the other hand, Chinese linguistic MA learners tend to avoid low probability modal verbs and instead prefer to use obligatory modal verb ‘should’, which expresses a relatively high degree of certainty. This finding is consistent with the results reported by K. Hyland & J. Milton (1997). This difference in modal verb usage may stem from learners’ inadequate linguistic knowledge, incomplete grasp of the interpersonal role of modal verb combinations, or limited understanding of how to effectively convey certainty and uncertainty in propositions.

In summary, Chinese linguistic EFL learners exhibit similarities with linguistic experts in their use of speculative verbs but also show significant differences, particularly in frequency and specific choices. They use hedging strategies such as impersonal constructions and modal verbs like ‘may’ and ‘might’ to achieve scientific objectivity in their writing. However, they tend to rely more on certain constructions influenced by Chinese academic norms, indicating their awareness of objectivity but also their novice status in academic writing. The key findings can be outlined as follows: (1) Chinese linguistic EFL learners resemble linguistic experts in the diversity of speculative verbs but differ



significantly in frequency and specific choices; (2) Chinese learners use hedging strategies like impersonal constructions to distance themselves from viewpoints and maintain objectivity. (3) Personal pronouns are commonly used by learners for literature review, while experts use shell nouns for concise summaries; (4) Chinese academic norms influence learners to avoid first-person expressions and use phrases like ‘the author suggests’ instead of ‘I suggest’. (5) Experts utilize probability modal verbs like ‘may’ and ‘might’ to express uncertainty, whereas learners tend to use ‘should’ with a higher degree of certainty. (6) The difference in modal verb usage may stem from learners’ insufficient linguistic knowledge or limited understanding of expressing certainty and doubt.

### 3.3.2. An investigation of quotative verbs

Quotative verbs were retrieved in AntConc, and their raw frequency and TPM were obtained for each corpus. Following this, the Log-likelihood and Chi-square Calculator 1.0 were utilized to pinpoint unique frequency characteristics for each quotative verb in CLMA\_C and ILJA\_C, and the outcomes are presented below.

**Table 10. Frequency distribution of quotative verbs across two corpora**

Quotative verbs	CLMA_C		ILJA_C		Log-likelihood	Significance( <i>p</i> )
	RF	TPM	RF	TPM		
claim	113	144.34	52	66.59	22.94	0.000
argue	97	123.90	168	215.14	-19.44	0.000
propose	96	122.63	59	75.55	8.82	0.003
suggest	84	107.30	121	154.95	-6.81	0.009
believe	68	86.86	17	21.77	32.64	0.000
consider	53	67.70	17	21.77	19.34	0.000

conclude	39	49.82	19	24.33	6.99	0.008
think	19	24.27	11	14.09	2.14	0.144
maintain	12	15.33	4	5.12	4.17	0.041
indicate	10	12.77	11	14.09	-0.05	0.823
contend	2	2.55	9	11.53	-4.84	0.028
allege	1	1.28	2	2.56	-0.34	0.558
TOTAL	594	758.75	490	627.48	9.73	0.002

Table 10 indicates that Chinese linguistic EFL learners and experts have some similarities. Specifically, they use a wide range of quotative verbs in their writing instead of relying on a few specific ones. There are similarities in the proportion of some high-frequency quotative verbs, such as ‘claim’ (19.02% in CLMA\_C and 10.61% in ILJA\_C) and ‘propose’ (16.16% in CLMA\_C and 12.04% in ILJA\_C). These findings suggest that Chinese linguistic EFL learners are capable of using a variety of quotative verbs in a manner similar to that of linguistic experts. In examining the distinctions between the two corpora, this study revealed notable variations in the frequency of quotative verbs, indicating that CLMA\_C demonstrates a higher frequency compared to ILJA\_C ( $p=0.002 < 0.05$ ). Further examination uncovered distinct variations in the utilization of these verbs between Chinese EFL learners and experts in linguistics. Specifically, novices used ‘claim’ most frequently (19.02%, 113 out of 594), while experts preferred ‘argue’ (34.29%, 168 out of 490). Additionally, novices significantly used ‘argue’ and ‘suggest’ less frequently than experts ( $LL=-19.44, p<0.05$ ;  $LL=-6.81, p<0.05$ ), but used ‘believe’, ‘claim’ and ‘consider’ more frequently ( $LL=32.64, p<0.05$ ;  $LL=22.94, p<0.05$ ;  $LL=19.34, p<0.05$ ). These findings suggest that while Chinese EFL learners are capable of using a variety of quotative verbs like experts, there are notable differences in their specific choice

of verbs. To gain a deeper understanding of how Chinese MA novices and experts use quotative verbs to achieve their rhetorical functions of hedging, it is important to examine their usage patterns in specific contexts and identify the specific collocational features associated with them. Quotative verbs, also called ‘reporting verbs’ by M. Charles (2006), are primarily used to convey information from other sources. In academic discourse, they are often used to report the views or findings of other researchers and are an important hedging rhetorical device for citation [Hyland 2005]. In terms of collocation, quotative verbs typically follow ‘that’ in R1, forming a clause introduced by ‘that’. Therefore, the criteria for determining quotative verbs are: the information that is quoted is clearly introduced by a quotative verb, the source from which the quote originates is distinctly specified within the context, and the quotative verb typically follows ‘that’ in R1. Therefore, we use ‘quotative verb + that’ as the search pattern and manually identify the concordance lines that contain this pattern according to the above criteria. In this section, we take ‘claim’ as example because it is high-frequency quotative verb and the difference of it is quite clear in the two corpora.

**Table 11. Collocational information of ‘claim that’ in CLMA\_C and ILJA\_C**

CLMA_C				ILJA_C			
Collocate	FreqL	FreqR	Likelihood	Collocate	FreqL	FreqR	Likelihood
)	37	0	285.293	it	9	5	81.307
(	37	0	238.285	is	3	10	51.306
the	5	40	163.423	)	9	0	42.606
is	4	43	141.682	(	9	0	39.738
of	17	18	116.632	to	6	6	37.632
and	22	5	85.989	has	6	1	36.191

a	5	16	85.525	been	6	0	29.830
schema	0	9	55.919	can	0	5	23.646
language	2	8	39.811	as	4	7	18.142
discourse	5	4	35.497	by	0	5	12.609

Considering the limited space available, we have included only the most significant collocates of the search item ‘claim that’ in this section. These collocates are ranked based on their likelihood scores in Table 11, and we have listed the top 10. The significant collocates of ‘claim that’ in CLMA\_C are ‘)’, ‘(’, ‘the’, ‘is’, ‘of’, ‘and’, ‘a’, ‘schema’, ‘language’ and ‘discourse’. While the significant collocates of ‘claim that’ in ILJA\_C are shown differently, i.e., ‘it’, ‘is’, ‘)’, ‘(’, ‘to’, ‘has’, ‘been’, ‘can’, ‘as’ and ‘by’. Table 11 reveals both similarities and notable differences in the collocational patterns of ‘claim that’ used by novices and experts.

In terms of similarities, both novices and experts frequently utilize the collocates ‘)’ and ‘(’. According to Table 11, the collocational strengths of ‘)’ and ‘(’ in the CLMA\_C and ILJA\_C corpora are ‘285.293; 238.285’ and ‘42.606; 39.738’ respectively (see Table 11). These collocations are among the most frequently used in both corpora. They form the primary linguistic manifestation of the quotative verb ‘claim’: ‘surname (year) + claim/claims/claimed + that’, as noted by K. Hyland (1998), F. Xu (2014), and B. Lou (2020). In this structure, ‘surname’ typically represents the experts of the reported information, such as an authoritative figure or well-known scholar. By engaging in this practice, the author fosters connections with various academic communities, improves group flexibility, and reinforces their standing in the academic sphere, thereby bolstering the scientific validity and trustworthiness of the research outcomes. For example:

(24) “Ungerer and Schmid (1996) claim that image schema is the basic cognitive mode which comes from the process of our everyday experience interacts with the world.” (CLMA\_C)

(25) “De Haan (2001), on the other hand, claims that English ‘must’ is always epistemic, contrary to Dutch ‘moeten’, which can have both an epistemic and an evidential reading.” (ILJA\_C)

In sentences (24) and (25), this structure serves several purposes. Firstly, it clearly indicates the views or claims of the quoted experts and provides precise references to the sources. Secondly, the author’s use of secondary evidence is left hedged, implying that the author is attempting to avoid overly subjective expressions and to achieve a more objective and accurate expression. Additionally, in academic writing, employing this structure enhances the credibility and authority of the article, as it demonstrates the author’s rigorous approach and standards when evaluating and analyzing previous research. In both sentences, ‘Ungerer’, ‘Schmid’, and ‘De Haan’ are recognized figures in the realm of linguistics, and their cited views help avoid overly subjective discourse, ensuring the authoritative nature of the argument and contributing to the author’s scholarly identity. Moreover, ‘(1996)’ and ‘(2001)’ provide readers with the sources and background of relevant studies on image schema and the modal verb ‘must’ in linguistics, facilitating the sharing of academic background and knowledge dissemination. Based on this analysis, we can conclude that novices are capable of using the ‘surname (year) + claim/claims/claimed + that’ structure to quote authoritative information, supporting the author’s stance and argument.

In terms of differences, firstly, although novices are capable of using the ‘surname(year) + claim/claims/claimed + that’ structure to quote authoritative information, they have not yet realized that in English culture, the word ‘claim’ is often associated with conjunctions or phrases such as ‘on the other hand’, ‘however’ and ‘actually’, which indicate contrast or different opinions. These collocations are used to express negation, doubt, or uncertainty regarding the

reported information [Hunston 1995; Charles 2006]. When they want to express negation or negative viewpoints about others' ideas, they usually employ various hedging rhetorical strategies rather than relying solely on reporting verbs to soften their expression [Hyland 2002]. For example, in sentence (25), hedging rhetorical strategies were used by using 'on the other hand' and 'be...contrary to' to indicate that the author's viewpoint is not exclusive and that other possibilities exist. Specifically, the author cited the research findings of 'De Haan (2001)', claiming that 'must' in English always conveys epistemic meaning, while 'moeten' in Dutch can express both 'epistemic and evidential meanings'. These rhetorical strategies make the author's stance more objective and cautious, while also making it easier for readers to accept and understand the author's viewpoint.

Secondly, the strongest collocate in the ILJA\_C corpus is 'it' (Likelihood > 81.307) (see Table 11). Through concordance line observation and categorization, we can observe that 'it' forms another linguistic realization structure together with reporting verb 'claim': 'it is/was/has been + claimed + that'. Using the structure 'It has been claimed that' to report information is necessary in academic papers when we need to cite other researchers' viewpoints or research findings. This structure indicates that we have not personally conducted the research or experiments but are referencing the research outcomes of other scholars. It serves as a hedging strategy that helps maintain objectivity while citing others' viewpoints, avoiding overly subjective expressions. Simultaneously, this structure highlights the methodology employed in our research, which involves supporting our own research perspectives by referencing other scholars' findings. Therefore, in academic papers, using the 'It has been claimed that' structure to report information is a common practice. For example,

(26) "It has been claimed that computer mediated asynchronous text-based conferencing is a useful medium for developing argumentation skills (Andriessen, Baker, & Suthers, 2003)." (ILJA\_C)

This sentence uses the ‘It has been claimed that’ structure to cite the research findings of other scholars and presents a viewpoint that asynchronous text-based conferences on computers are a useful medium for developing argumentation skills. The source of this viewpoint is indicated as ‘(Andriessen, Baker, & Suthers, 2003)’. This citation method helps the author reference the views of other scholars in the paper and also indicates that the author’s own viewpoint is based on the research findings of other scholars. This citation approach enhances the objectivity of the paper, avoiding overly subjective expressions and thereby increasing the paper’s credibility.

In addition, experts may also use the structure ‘as claimed by + surname (year)’ to report information (see Table 11). Both ‘as claimed by +surname(year)’ and ‘it is/was/has been + claimed + that’ are utilized to reference the perspectives or research findings of other scholars and serve hedging purposes, but their grammar structures and rhetorical purposes slightly differ. The ‘as claimed by +surname (year)’ structure is usually used to cite the views or research findings of a specific scholar, emphasizing the source of the citation and the author’s viewpoint. This structure is often used to support the author’s own viewpoint or research findings. While the ‘it is/was/has been + claimed + that’ structure is more general and is commonly used to cite a widely accepted viewpoint or research findings, rather than attributing it to a specific scholar. This structure is employed to signify that a particular viewpoint or research finding has been widely acknowledged.

However, novices seldom use the above two structures and tend to excessively rely on the ‘surname (year) + claim/claims/claimed + that’ structure for academic citations. This indicates that novices lack diversity in the specific linguistic manifestations of hedging through quotative verbs. The reasons for this might be threefold: firstly, novices lack confidence in their own English academic writing abilities and excessively rely on authoritative voices to present their viewpoints, resulting in writing that may not be as sophisticated as experts;

secondly, novices may be unfamiliar with the evaluative meanings and rhetorical functions of quotative verbs [Bloch 2010]; thirdly, Chinese language emphasizes implicature and exhibits implicit discourse markers. Chinese EFL Learners are deeply influenced by Chinese thinking and have not fully acquired the explicit discourse marker function of English [Yang 2019].

In summary, Chinese EFL learners and experts alike demonstrate both commonalities and distinctions in their utilization of quotative verbs. Both learners and experts displayed a diverse range of quotative verbs, with notable similarities in the proportion of certain high-frequency verbs such as ‘claim’ and ‘propose’. However, significant differences were observed in the frequency and specific choices of quotative verbs between the two groups, with learners exhibiting a higher overall frequency. Further analysis revealed individual differences in verb usage, with novices favoring ‘claim’ while experts preferred ‘argue’. Despite their ability to use a variety of quotative verbs, learners exhibited differences in specific verb choices compared to experts.

Additionally, the collocational patterns of the verb ‘claim that’ were examined in both learner and expert corpora. While some similarities were noted, such as the frequent use of parentheses, differences in collocational patterns were also observed. Novices tended to rely on a specific structure for quoting authoritative information, while experts employed additional linguistic strategies, such as conjunctions and phrases indicating contrast or differing opinions. Moreover, experts frequently used the structure ‘It has been claimed that’ to cite research findings, enhancing objectivity and credibility in academic writing.

Overall, while novices demonstrated proficiency in using quotative verbs, their usage differed from that of experts, indicating potential areas for improvement in academic writing skills. These differences may be attributed to factors such as confidence levels, familiarity with evaluative meanings, and cultural influences on discourse markers.



### 3.3.3. An investigation of sensorial verbs

This section will firstly use AntConc to extract sensorial verbs from two different corpora and calculated their raw frequency and TPM. Next, Log-likelihood and Chi-square Calculator 1.0 will then be employed to compare the distinctive frequency features for each sensorial verb in CLMA\_C and ILJA\_C, and the results are presented below.

**Table 12. Frequency distribution of sensorial verbs across two corpora**

Sensorial verbs	CLMA_C		ILJA_C		Log-likelihood	Significance( <i>p</i> )
	RF	TPM	RF	TPM		
seem	542	692.32	273	349.60	89.80	0.000
appear	155	197.99	37	47.38	77.66	0.000
feel	42	53.65	31	39.70	1.64	0.201
sound	14	17.88	11	14.09	0.35	0.552
look	9	11.50	8	10.24	0.06	0.812
TOTAL	762	973.34	360	461.01	146.27	0.000

Table 12 shows that both groups exhibit similarities, as they use a diverse range of sensorial verbs instead of relying on a few specific ones. This suggests that novices also have the ability to employ a variety of sensorial verbs in a manner similar to linguistic experts. Regarding to the individual similarity, both novices and experts used ‘seem’ most frequently with the proportion of 71.13% (542 out of 762) and 75.83% (273 out of 360). The ranking of the proportion of other sensorial verbs is as follows: ‘appear’ (20.34%, 155 out of 762; 10.28%, 36 out of 360); ‘feel’ (5.51%, 42 out of 762; 8.61%, 31 out of 360); ‘sound’ (1.84%, 14 out of 762; 3.06%, 11 out of 360); ‘look’ (1.18%, 9 out of 762; 2.22%, 8 out of 360).

Regarding the disparities between the two corpora, this investigation identified notable disparities in the frequency of sensorial verbs, with CLMA\_C exhibiting a notably higher occurrence than ILJA\_C ( $p= 0.000 < 0.05$ ). Subsequent analysis revealed discernible disparities in the utilization of these verbs between Chinese linguistic EFL learners and experts. Specifically, novices exhibited significantly greater usage of ‘seem’ and ‘appear’ compared to experts (LL=89.80,  $p=0.00 <0.05$ ; LL=77.66,  $p=0.00 <0.05$ ). These findings indicate that, although Chinese EFL learners demonstrate proficiency in employing a diverse array of sensorial verbs similar to experts, discernible discrepancies exist in their specific verb selection. In academic writing, sensorial verbs serve to depict researchers’ observations, perceptions, or interpretations with objectivity and precision while acknowledging inherent study limitations and uncertainties using hedging strategies [Vass 2017]. To gain a more profound comprehension of how Chinese MA novices and experts utilize sensorial verbs to achieve hedging rhetorical functions, an in-depth examination of their usage patterns in specific contexts and identification of associated collocational features becomes imperative. For illustrative purposes in this section, we use ‘seem’ as it stands out as the sensorial verb most frequently utilized, highlighting clear differences between the two corpora.

**Table 13. Collocational information of ‘seem’ in CLMA\_C and ILJA\_C**

CLMA_C				ILJA_C			
Collocate	FreqL	FreqR	Likelihood	Collocate	FreqL	FreqR	Likelihood
it	119	8	292.493	to	22	460	525.371
to	14	174	159.368	be	6	137	217.092
that	23	85	132.247	it	119	12	191.239
be	3	71	101.325	would	35	3	77.895
quite	2	7	30.012	that	50	97	68.811

less	0	9	21.631	have	5	52	55.217
have	2	22	19.057	reasonable	0	8	48.627
justifiable	0	2	18.712	of	81	42	47.414
odd	0	2	18.712	contradictory	2	3	27.246
learners	12	6	18.275	less	2	15	25.368
more	2	21	17.027	indicate	1	10	24.483
regard	2	3	17.500	more	7	34	22.751
category	4	1	15.039	this	49	22	20.268
very	3	5	14.387	problematic	3	3	17.812
although	7	0	12.666	there	18	3	17.24
true	1	3	12.424	suggest	0	9	15.675
difficult	0	5	12.006	likely	1	8	13.252

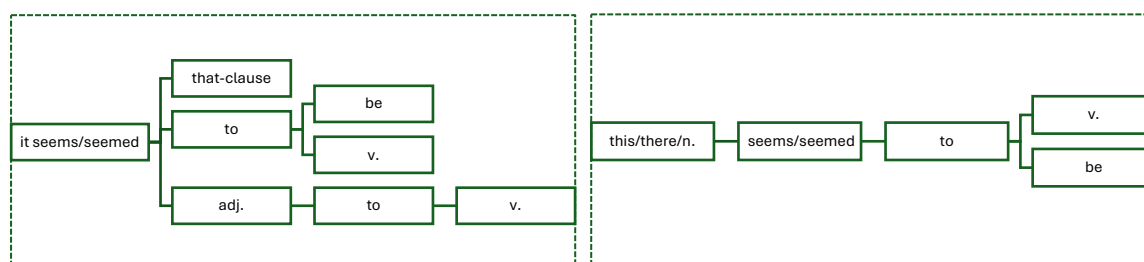
Considering the limited space available, we have included only the most significant collocates of the search item ‘seem’ in this section. In CLMA\_C, the significant collocates of ‘seem’ are ‘it’, ‘to’, ‘that’, ‘be’, ‘quite’, ‘less’, ‘have’, ‘justifiable’, ‘odd’, ‘learners’, ‘more’, ‘regard’, ‘category’, ‘very’, ‘although’, ‘true’, ‘difficult’. In contrast, the significant collocates of ‘seem’ in ILJA\_C include ‘to’, ‘be’, ‘it’, ‘would’, ‘that’, ‘have’, ‘reasonable’, ‘of’, ‘contradictory’, ‘less’, ‘indicate’, ‘more’, ‘this’, ‘problematic’, ‘there’, ‘suggest’, ‘likely’. Table 13 displays similarities and notable variations in how both Chinese EFL learners and experts employ collocations involving the verb ‘seem’.

Firstly, in terms of grammatical collocations, specifically in the aspect of ‘colligation’ proposed by J. Sinclair (2004), novices demonstrate a certain degree of similarity with experts in using grammatical collocation structures related to ‘seem’. These structures can be broadly categorized into two main types: ‘it

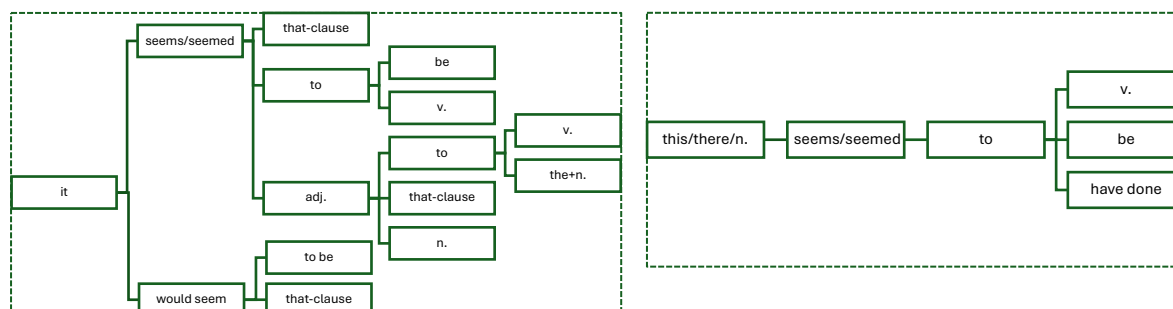
seems/seemed + that-clause/to (be/v.) /adj.’ and ‘this/there/n. + seems/seemed + to be/to do’.

The first category involves the use of ‘it’ as the formal subject, collocating with ‘seems/seemed’ in the third person singular or past tense, leading to three main patterns: the first is ‘it seems/seemed that-clause’; the second is ‘it seems/seemed + to be/v.’; and the third is ‘it seems/seemed + adj. + to v.’.

The second category involves the use of ‘this/there/n.’ collocating with ‘seems’ in the third person singular or past tense, leading to ‘to be/v.’ structure. The specific collocation patterns are illustrated in Figure 3 and Figure 4.



**Figure 3. Colligational information of ‘seem’ in CLMA\_C**



**Figure 4. Colligational information of ‘seem’ in ILJA\_C**

However, upon meticulous examination, it became evident that the colligational structure ‘it would seem to be/that-clause’ was absent from the CLMA\_C corpus (see Figure 4). This may be attributed to novices having fewer varied colligations for ‘seem’ compared to experts, leading to a limited diversity in grammatical collocations. Additionally, novices may not have fully acquired the hedging pragmatic function of the colligation structure ‘it would seem to be/that-clause’. Combinations of hedging expressions can possess specific pragmatic functions. For example, the ‘it would seem to be/that-clause’ pattern

(see Figure 5) includes the volitional modal verb ‘would’ and the sensorial verb ‘seem’, forming a combination with potential hedging pragmatic function. Experts often use this lexical pattern in academic writing to soften the tone of the statement and show respect for the reader’s perspective. It is a way of presenting an idea without sounding overly confident or imposing one’s beliefs on the reader.

(27) “For more advanced language learners, it would seem that a comprehensive approach might also be effective if they are able to attend to a range of linguistic foci.” (ILJA\_C)

In (27), by using ‘It would seem that...’, the author softens the statement and acknowledges that their suggestion or observation may not be definitive. Instead, it is presented as a possibility or hypothesis, leaving room for other viewpoints or interpretations. The usage of ‘It would seem that...’ indicates that, given ‘the available evidence’ or ‘logical reasoning’, the holistic approach could potentially benefit advanced language learners, although the author refrains from making definitive or absolute assertions. This strategy is particularly useful in academic writing where the emphasis is on presenting evidence, discussing findings, and maintaining a scholarly tone. By using ‘It would seem that...’, the author is inviting the reader to consider the proposition rather than imposing it as an absolute truth. It demonstrates humility and open-mindedness, which are essential traits in academic discourse. Overall, the use of ‘It would seem that...’ in the example aligns with the purpose of expressing politeness and caution while presenting an idea in academic writing. It helps to create a balanced and respectful tone, fostering a constructive and thoughtful dialogue with the reader.

Besides, experienced experts preferred adding ‘seem to’ in front of ‘suggest/indicate’, forming a pattern ‘... seem/seems/seemed to suggest/indicate that clause’. For example,

(28) “Our quantitative analysis seems to suggest that there may be negative effects when changing the lecturing language to English in terms of time taken and speed of delivery.” (ILJA\_C)

Experts often use the phrase ‘seem to’ in front of ‘suggest’ or ‘indicate’ as an additional hedging strategy to express a sense of caution or uncertainty. While ‘suggest’ or ‘indicate’ themselves can be considered hedging verbs as they convey possibilities rather than absolute conclusions, the addition of ‘seem to’ further emphasizes that the evidence or data only appears to point in a particular direction, and there might be other interpretations or factors to consider.

In (28), the author is discussing the results of a quantitative analysis on the impact of changing the language of lecturing to English. By using ‘seems to suggest’, the author is being cautious about the findings and not presenting them as definitive conclusions. The phrase ‘seems to suggest’ indicates that the analysis provides evidence that points towards the presence of ‘negative effects’, but it does not firmly establish this as an absolute fact. The addition of ‘seems to’ serves to soften the assertion and acknowledge the possibility of other factors that may influence the results. It aligns with the scholarly practice of hedging in academic writing, where authors often avoid making absolute claims and acknowledge the limitations or uncertainties in their research.

Overall, the use of ‘seem to’ ahead of ‘suggest/indicate’ in the example exemplifies the author’s careful approach to presenting research findings, showing respect for potential alternative interpretations and encouraging further exploration of the topic. It also contributes to the overall academic tone and credibility of the writing.

In summary, this research explores how sensorial verbs are similarly and differently used by Chinese EFL learners and experts. Both groups demonstrate a diverse range of sensorial verbs, suggesting that novices possess the ability to utilize these verbs similarly to experts. Notably, both novices and experts most frequently used the verb ‘seem’. However, significant differences were observed

in the frequency of sensorial verbs between the two corpora, with novices exhibiting a higher occurrence in CLMA\_C compared to ILJA\_C. Subsequent analysis revealed their disparities, with novices using ‘seem’ and ‘appear’ more frequently than experts.

Regarding grammatical collocations, novices displayed some similarities with experts in using certain structures related to the verb ‘seem’. However, novices showed a lower richness of colligations compared to experts, indicating a lack of diversity in grammatical collocations. For instance, the colligational structure ‘it would seem to be/that-clause’ was not present in the CLMA\_C corpus, suggesting that novices may not have fully acquired the hedging pragmatic function associated with this structure.

In contrast, experts frequently employed hedging strategies such as ‘It would seem that...’ and ‘seem to suggest/indicate’, which serve to soften statements and acknowledge uncertainties in academic writing. These strategies contribute to the overall scholarly tone and credibility of the writing by demonstrating humility and open-mindedness.

Overall, while novices demonstrate proficiency in using sensorial verbs, differences exist in their specific verb selection and employment of hedging strategies compared to experts. Further investigation into their usage patterns and associated collocational features is necessary to gain a deeper understanding of how these verbs are utilized to achieve rhetorical functions in academic writing.

### **3.4. Constructing authorial identity in academic writing: An investigation of self-mentions**

Interplay between language and identity has been a central focus of sociolinguistic inquiry for some time, gaining particular significance in recent years as identity is increasingly recognized as something actively achieved through our interactions with others [Benwell and Stokoe 2006]. Identity is

understood as an individual's connection to their social environment, a collaborative, two-way process. Language facilitates the construction and presentation of a coherent self to others by embedding us within webs of shared understandings, interests, and meanings. Our identities, both present and potential, are shaped through participation and intertwined with contexts, relationships, and the rhetorical strategies we employ in everyday interactions. Consequently, our use of communal discourses serves as a means of asserting or challenging our membership in social groups and defining ourselves in relation to others. Identity thus serves to delineate both our commonalities and differences, and for academics, it is instrumental in establishing credibility within their fields and cultivating individual reputations.

Crafting a self-representation amid the standardizing conventions of disciplinary discourses demands skill, entailing both the acknowledgment and strategic utilization of communal constraints. However, this endeavour poses a particular challenge for EFL learners. Embracing the notion that identity is shaped through discourse necessitates a method for examining how individuals routinely construct markers of their identity through interaction. This section delves into the construction of authorial identity through a corpus-based analysis of how Chinese EFL learners and experts employ two metadiscourse markers: self-mention markers and engagement markers. By examining the nuanced characteristics of these markers, the study aims to identify the distinctive patterns employed by each group and elucidate their implications for the construction of authorial identity.

Academic discourse serves as a crucial medium for knowledge dissemination and scholarly exchange. As functional linguistics and the sociology of scientific knowledge have evolved, there is growing recognition that academic discourse serves not only to convey scientific information and generate credible texts but also to articulate intricate interpersonal implications [Hyland 2002; Jiang 2017; Jiang 2020]. The credibility and acceptance of a paper depend not



only on the reliability and validity of itself but also on the persuasiveness of the author's argumentation, namely academic persuasion. The premise is that the presentation of viewpoints should adhere to academic discourse conventions to resonate with readers [Jiang 2016]. The process of persuasion and argumentation is most directly manifested when authors intervene in the discourse explicitly through linguistic means, organize text segments, evaluate discourse content, and guide readers in co-constructing discourse [Jiang 2019]. For example, the phrase 'we show that' contrasts with 'the results show that' highlights the author's involvement in the research findings, aiding readers in identifying the author's innovative contribution. Thus, self-mention is instrumental in fostering interactive persuasion and elevating the author's prominence within academic discourse.

Self-mentions, such as 'I', 'my' and 'we', signal the author's presence and fostering engagement with the reader in academic discourse. They serve as essential tools for establishing authorial identity and advancing persuasive arguments. Despite their growing importance in academic writing and knowledge construction, Chinese novice writers face challenges in utilizing self-mentions effectively to craft their academic texts and present coherent arguments [Wang and Jiang 2019]. This difficulty is compounded by the limited emphasis on the lexical and grammatical aspects of self-mention in existing literature, which tends to focus more on their rhetorical functions [Hyland 2001; Jiang 2020; Walková 2019]. Furthermore, traditional Chinese pedagogy often neglects the instruction of self-mentions, with some educational materials even discouraging their academic employment, further hindering students' acquisition of effective self-referential techniques [Bennett 2009; Jiang and Hyland 2020]. M. Walková (2019) suggests that rhetorical function is just one aspect of self-mention usage, and phrase structure pose greater obstacles for second language authors to master self-mention.

Thus, by employing self-compiled corpora from Chinese linguistic MA theses and international journal articles, this research analyzes the collocational and chunk patterns of ‘we’ used by novice writers versus experts. The aim is to offer insights and references for improving academic English writing instruction through comparative analysis.

Consequently, this study primarily focuses on three tasks concerning the examination of self-mention ‘we’: firstly, it extracts and categorizes strong left and right collocates of ‘we’ from the standpoints of disciplinary characteristics and semantic function. Secondly, it employs corpus-driven techniques to calculate high-frequency chunks preceded by ‘we’ in both corpora and summarizes the general chunk characteristics in CLMA\_C and ILJA\_C; and finally, conducting a comparative analysis of the collocation features and chunk characteristics between Chinese EFL learners and international journal experts.

Based on these three tasks, the research methodology of this study primarily employs a corpus-driven approach and academic discourse analysis. The examination of chunk characteristics of self-mention ‘we’ necessitates a bottom-up corpus-driven approach, while the comparison of collocation features of self-mention ‘we’ between Chinese linguistic EFL learners and experts requires detailed analysis. Understanding the distinctions amid the two groups relies on the application of academic discourse analysis.

Earlier studies have recognized multiple roles of self-mention in academic persuasion. K. Hyland (2002) found that self-mention ‘we’ enhances the persuasiveness of academic discourse through five main functions: stating research objectives, introducing research processes, explaining arguments, presenting research results, and expressing personal contributions. R. Tang & S. John (1999) argued that the persuasive function of self-mention ‘we’ constructs different academic identities for authors, such as ‘guide’, ‘recounting’, and ‘opinion-holder’, thereby influencing readers’ acceptance of viewpoints. Additionally, disciplinary differences and author groups can also influence self-

mentions' employment. K. Hyland (2001) and F. Jiang (2017) found that self-mention 'we' is significantly less frequent in natural sciences compared to humanities and social sciences. Nonetheless, recent research reveals a considerable shift in this pattern, highlighting a significant rise in the prevalence of self-mention in the natural sciences, mainly through the use of 'we'. [Jiang and Hyland 2020]. Regarding differences in the author groups using self-mention, M. Walková (2019) found that second language learners tend to use self-mention 'we' less frequently compared to native speakers, thereby downplaying their personal knowledge contributions. K. Fløttum (2007) suggests that this covert presentation contradicts the English writing culture, which emphasizes explicit presentation of key information. J. Wang & F. Jiang (2019) found instances of underuse and misuse of 'we' in Chinese students' academic writing, suggesting that differences in self-mention usage between students and experts warrant further systematic investigation.

The majority of previous studies in this area have primarily focused on identifying various rhetorical functions of self-mention 'we', with minimal attention paid to collocational phrase structures. In fact, for second language (L2) writers, mastery of rhetorical functions represents an advanced writing skill, which cannot be achieved without a proficient command of the structural aspects of self-mention phrases.

Given this context, the present study employs a corpus-driven approach to contrast the collocational and chunk patterns of the self-mention pronoun 'we' between Chinese linguistic EFL learners and linguistic experts. Therefore, the research questions are as follows: (1) What are the collocation characteristics and chunk features of self-mention 'we' in Chinese linguistic MA theses and international linguistic journal articles? (2) Concerning these collocation characteristics and chunk features, what similarities and distinctions do they display, and what factors contribute to these differences?

In addressing the two research inquiries, the following section will first examine the collocational attributes of self-mention ‘we’ in CLMA\_C and ILJA\_C. It will then proceed to analyze their similarities, differences, and the reasons behind these findings. Following the same methodology, chunk features of self-mention ‘we’ will then be examined.

Firstly, overall collocation characteristics of ‘we’ in CLMA\_C and ILJA\_C are display here.

LancsBox 6.0 [Brezina 2020] was utilized to calculate the frequency of left and right collocates of ‘we’ in the CLMA\_C and ILJA\_C corpora respectively. These frequencies were then sorted by MI3 score, a metric designed to rebalance the Mutual Information (MI) score by assigning greater importance to frequent words and lesser importance to infrequent ones [Oakes 1998]. This approach is influential for identifying strong collocates of a given search item. The detailed data concerning the left collocates of ‘we’ are presented in Table 14.

**Table 14. Left collocates of ‘we’ in CLMA\_C and ILJA\_C**

CLMA_C			ILJA_C		
Collocates	MI3	Frequency	Collocates	MI3	Frequency
in	21.70	78	in	22.20	671
of	20.63	569	of	21.40	618
this	19.97	511	this	20.67	297
from	19.96	236	and	20.63	483
and	19.51	206	as	19.77	278
above	19.29	358	what	18.08	104
as	18.74	95	if	17.71	79
so	18.05	201	is	17.59	171
is	17.62	92	for	17.47	149

when	17.48	186	here	16.51	50
------	-------	-----	------	-------	----

Notes: To facilitate the presentation and comparison of data analysis, the minimum threshold for MI3 is set to 16, and the span is set from left 5 to right 5.

Regarding the similarities, the words ‘in’, ‘of’, ‘this’, ‘and’, ‘as’ and ‘is’ are observed as collocates of ‘we’. This indicates that these words are commonly used in association with ‘we’, whether by Chinese master’s students learning English or by authors of international authoritative journals. These words are primarily utilized for constructing discourse structures and logic, as exemplified by phrases such as ‘in this section, we...’, ‘as we have seen...’ and ‘is what we call...’.

In terms of their differences, four words appear exclusively in each set of data: ‘from’, ‘above’, ‘so’ and ‘when’ in CLMA\_C, and ‘what’, ‘if’, ‘for’ and ‘here’ in ILJA\_C. These words reflect different linguistic functions and styles. ‘From’, ‘above’, ‘so’ and ‘when’ are predominantly utilized for depicting relationships such as source, context, causality, and time, as seen in phrases like ‘from this perspective, we...’, ‘above all, we...’, ‘so we can conclude...’ and ‘when we...’, which are exemplified in complete sentences (30) to (33). These collocations align with the hypothesis that Chinese master’s students emphasize guiding readers through discourse structure and logic when using ‘we’.

(29) “From this perspective, we should interpret the features of human communications and social processes, symbolization of social structures, and the language process in which social members constitute social semiotics.” (CLMA\_C).

Sentence (29) elucidates the purpose, which is to understand social semiotics, and subsequently delineates the steps that ‘we’ need to undertake.

(30) “From above all we can conclude that conjunctive Adjuncts are not the constituents within Mood structure but they form a constituent on their own, which is the part of the clause.” (CLMA\_C).

Sentence (30) highlights the shared journey of analysis ‘we’ have undertaken and presents the conclusion.

(31) “So we can conclude that the source selection in the legal metaphors, to some extent, is subject to the already existing targets.” (CLMA\_C).

Sentence (31) emphasizes the shared investigation ‘we’ are conducting and the resulting observation.

(32) “However, when we move onto the next stage of Interpretation, we find that the authentic recording of the direct citing form definite news source is rather limited.” (CLMA\_C).

Sentence (32) highlights the transition in the discussion ‘we’ are making and presents a challenge encountered.

In all four sentences, the use of ‘we’ creates a sense of a shared journey with the reader, guiding the reader through their analysis or interpretation, making the thought process and reasoning clear. This reinforces the hypothesis that Chinese EFL learners might use ‘we’ strategically to enhance clarity and navigate readers by the logic of their arguments.

While ‘what’, ‘if’, ‘for’ and ‘here’ are mainly used to express questions, hypotheses, purposes, and positions, as shown in phrases like ‘what we propose...’, ‘if we assume...’, ‘for this reason, we...’ and ‘here we present...’. These collocations also align with the hypothesis that disciplinary experts often give precedence to the utilization of ‘we’ to articulate their academic viewpoints and innovations.

In academic discourse, the strategic use of terms like ‘what’, ‘if’, ‘for’, and ‘here’ alongside ‘we’ is essential. These words help convey queries, hypotheses, objectives, and stances, thereby demonstrating the academic perspectives and advancements of authors in leading international journals or disciplinary experts. The following are examples and their explanations:

(33) “What we found was that Victoria was seemingly working through three central questions in her accounts of her L2 writing.” (ILJA\_C).

In this example, ‘what’ introduces the content discovered by the authors, emphasizing the three central questions that Victoria appears to be addressing in her description of second language writing.

(34) “If we assume that L2 speakers are much more likely than L1 speakers to face problems.” (ILJA\_C).

Here, ‘if’ introduces a hypothetical scenario where it is assumed that L2 speakers are much more likely than L1 speakers to encounter problems.

(35) “For this reason, we might expect to find it in the informal written conversations that occur in online environments.” (ILJA\_C).

In this instance, ‘for’ introduces the reason or justification behind an expectation, indicating that it is logical to anticipate finding something in informal written conversations that occur in online environments.

(36) “Here, we present a conceptual model for studying adolescent L2 writers and their writing and we identify distinct vantage points for understanding and researching this population.” (ILJA\_C).

In this instance, ‘here’ serves to denote the immediate setting or context of the discourse, underscoring the authors’ presentation of a conceptual framework for investigating adolescent L2 writers and their writing. It also delineates specific perspectives crucial for comprehending and researching this demographic.

These examples demonstrate how using these key words in conjunction with ‘we’ can emphasize authors’ academic positions, knowledge innovations, and research viewpoints. Such clear expression aids readers in understanding authors’ perspectives and facilitates comprehension and evaluation of research results.

In studies on the academic authorial identity construction, scholars frequently analyze the lexical elements that appear immediately after the pronoun ‘we’ to gain insight into its functional role in academic writing [Jiang and Wang 2021]. Thus, the calculated numerical data pertaining to the right collocates of

‘we’ in CLMA\_C and ILJA\_C are presented in Table 15, organized in descending order based on their MI3 scores.

**Table 15. Right collocates of ‘we’ in CLMA\_C and ILJA\_C**

CLMA_C			ILJA_C		
Collocates	MI3	Frequency	Collocates	MI3	Frequency
can	24.18	568	be	21.54	599
see	21.99	199	have	20.43	276
find	21.32	197	can	19.86	181
be	20.37	459	see	19.66	147
will	20.22	163	will	19.14	113
have	20.10	234	find	18.33	95
know	19.69	91	believe	18.08	51
say	17.90	69	do	18.03	115
get	17.04	53	discuss	17.42	62
should	17.03	66	know	17.32	65
need	16.66	43	use	17.08	100
mention	16.63	42	argue	16.88	52
discuss	16.48	44	would	16.64	62
do	16.33	66	examine	16.61	53
conclude	16.20	32	observe	16.58	40
must	16.08	34	need	16.40	51
use	16.06	80	consider	15.99	48
analyze	15.93	50	present	15.88	42



go	15.77	33	focus	15.41	40
think	15.49	37	identify	15.11	38

Notes: To facilitate the presentation and comparison of data analysis, the maximum threshold for MI3 is set to 15, and the span is set from left 5 to right 5.

Due to the predominance of verbs as right collocates in both corpora (exceeding 90%), we categorize them according to their semantic roles as verbs, following D. Biber (1999). The specific categorization is illustrated in the Table 16.

**Table 16. Semantic functions of the right collocates of ‘we’ in CLMA\_C and ILJA\_C**

Semantic functions	CLMA_C	ILJA_C
State or attribute	be, have	be, have
Ability or possibility	can, will, should, must	can, will, would, should
Cognition or judgment	see, find, know, say, conclude, think	see, find, know, believe, argue, consider,
Action or process	do, discuss, use, get, mention, analyze, go	do, discuss, use, examine, observe, present, focus, identify
Necessity or purpose	need	need

From Table 16, it is evident that both Chinese EFL learners and experts of international authoritative journals exhibit consistency in the classification of semantic functions. When paired with verbs, they generally reflect a scientific research status, demonstrating academic credibility and cognition, and describing academic research processes and purposes. Examples include phrases like ‘we have made a/an... analysis’, ‘we can assume’, ‘we see... as’ and ‘we discuss...’.

However, there are differences in the specific vocabulary choices for the semantic functions of ‘Cognition or judgment’ and ‘Action or process’. In the ‘Cognition or judgment’ category, Chinese EFL learners tend to use more

cognition-oriented verbs such as ‘say’ and ‘conclude’ to describe their findings or conclusions (see Table 16). For instance:

(37) “For example, when we say ‘by and large’, it only requires retrieving from the memory knowledge of the idiom.” (CLMA\_C)

(38) “We conclude that the reverse form ‘Ground-Figure’ in news headlines greatly contributes to the strong emphasis of certain information.” (CLMA\_C)

On the other hand, experts tend to use verbs from the ‘judgment’ category to engage in subjective arguments, such as ‘believe’, ‘argue’ and ‘consider’. These verbs, when paired with ‘we’, indicate the authors’ stance and guide the reader on interpreting the presented information. For example:

(39) “We believe that the observed pattern of use of boosters can be plausibly explained in terms of the nature of the supports that claims in empirical vs. non-empirical academic articles are typically based on.” (ILJA\_C)

Sometimes, they also utilize verbs from the ‘possibility’ category such as ‘would’ to convey a certain degree of humility and caution, as exemplified by:

(40) “We would argue that such a dialogue is important for at least two reasons.” (ILJA\_C)

Moreover, Chinese EFL learners tend to use more general action verbs (e.g., ‘do’, ‘get’) and ‘mention’ to introduce topics when paired with verbs from the ‘Action or process’ category, reflecting their developing research skills. However, they also employ verbs like ‘analyze’, indicating deeper analysis. Examples include:

(41) “Through analysis, we get that Chinese writers tend to employ more generic headings than that of Canadian writers.” (CLMA\_C)

(42) “The second point we mention here is the features of metaphor.” (CLMA\_C)

(43) “On the one hand, the interaction we discuss in this study is a sort of social interaction.” (CLMA\_C)

In contrast, experts incline to employ action verbs that reflect more rigorous research activities like examining, observing, presenting, and identifying. Their process verbs focus on collaboration (e.g., ‘discuss’) and applying methods to reach focused conclusions. Examples include:

(44) “In the second part, we examine themes that emerged across the various interviews.” (ILJA\_C)

(45) “We observe three discourse markers – you know, and, and right?” (ILJA\_C)

Secondly, the overall chunk characteristics of ‘we’ in CLMA\_C and ILJA\_C are then presented in the following part.

This section mainly presents the chunk characteristics of Chinese linguistic EFL learners and experts aided with LanksBox 6.0. Table 17 displays the four-word chunks in patterns guided by ‘we’.

**Table 17. Four-word chunks characteristics guided by ‘we’ in CLMA\_C and ILJA\_C**

CLMA_C			ILJA_C		
Frequency	Text span	4-word chunks	Frequency	Text span	4-word chunks
68	24	we can see that	12	10	We believe that the
28	14	we can find that	8	5	we were able to
26	13	we can see the	7	6	we would like to
22	9	we can say that	7	5	we would argue that
19	7	we find that the	7	6	we can see that
15	9	we can conclude that			
13	5	we are going to			

11	5	We can see from
10	7	we can get the
9	7	we can find out
8	6	we can find the
6	6	we know that the

Notes: The chunks appear with a minimum frequency of 6 times, and the text span of chunks appears with a minimum frequency of 5 times.

According to K. Hyland (2008), chunks can be functionally categorized into three types: research-oriented, text-oriented, and participant-oriented. Research-oriented chunks aid writers in structuring their engagements and encounters in practical contexts. Text-oriented chunks focus on organizing textual content and its significance as a communication or argument. Participant-oriented chunks serve to forge a connection between the writer and the reader, conveying the writer’s perspectives, assessments, and positions. Based on the three types of chunks above, Table 18 displays the categorized four-word chunks and their percentage according to the statistics in Table 17.

**Table 18. Functional categories of four-word chunks guided by ‘we’ in CLMA\_C and ILJA\_C**

Types	CLMA_C	Percentage	ILJA_C	Percentage
Research-oriented chunks	we can find that, we can find the, we can conclude that, we can get the, we can find out	33.33%		
Text-oriented chunks	we are going to, we can see from	11.43%		

---

Participant-oriented chunks	we can see that, we can see the, we can say that	55.24%	we can see that, we believe that the, we were able to, we would like to, we would argue that	100%
-----------------------------	--	--------	--	------

---

A thematic analysis of functional categories of four-word Chunks with ‘We’ in CLMA\_C (Chinese EFL learners) and ILJA\_C (expert writers) reveals three key themes: functional divergence and communicative priorities.

As for functional divergence, the cohorts exhibit marked differences in chunk usage. Chinese EFL learners heavily employ research-oriented (33.33%) and text-oriented chunks (11.43%) (e.g., “we can conclude that”), focusing on procedural explanations (Example 47). In contrast, experts exclusively use participant-oriented chunks (100%) (e.g., “We believe that the”) to convey evaluation and engage readers (Example 48). This divergence highlights learners’ emphasis on methodological transparency versus experts’ prioritization of scholarly dialogue.

(46) “By analyzing these processes, we can find that quotation marks are used in these processes for three times.” (CLMA\_C)

(47) “We would argue that such a dialogue is important for at least two reasons.” (ILJA\_C)

Concerning to communicative priorities, thematic patterns reflect distinct rhetorical objectives. Learners’ chunk usage aligns with data-centric exposition, framing writing as a process of factual reporting. Experts, however, deploy chunks to construct author-reader rapport and assert epistemic authority, reflecting deeper engagement with argumentation and disciplinary discourse. This aligns with studies showing Chinese learners’ underuse of evaluative metadiscourse in identity construction compared to experts (Sun 2020; Lou 2020).

## CHAPTER III CONCLUDING REMARKS

Chapter III delved into rhetorical hype, hedging strategies, and the construction of authorial identity using metadiscourse features (such as certainty stance adverbs, maximizers, hedges, self-mentions, and engagement markers) in Chinese linguistic EFL learners' and linguistic experts' academic texts.

Firstly, the analysis of certainty stance adverbs and maximizers uncovered notable differences in their application between Chinese EFL learners and expert writers. These findings highlight variations in syntactic structures and the learners' partial grasp of academic norms when employing rhetorical hype strategies. This observation aligns with K. Hyland and F. Jiang's (2021) study on metadiscourse features in academic discourse related to hype strategies.

Likewise, the analysis of lexical verb hedging in Chinese MA theses in linguistics revealed discrepancies in both frequency and collocational patterns between Chinese EFL learners and expert writers. These differences underscore the necessity for novice researchers to develop a more refined awareness of hedging strategies to achieve greater fluency in academic writing. This perspective aligns closely with K. Hyland's (1998) assertion regarding the importance of strengthening hedging awareness among EFL learners.

Furthermore, our analysis relating to self-mentions and engagement markers, particularly the usage of 'we', elucidated significant disparities between Chinese linguistic MA theses and international journal articles. While Chinese EFL learners predominantly employed 'we' for discourse construction, experts utilized it to express hypotheses and positions, indicating differing approaches to reader engagement and evaluation. This finding further corroborates K. Hyland's (2005) assertion that the pragmatic usage of the first-person pronoun 'we' in academic writing varies among English learners from different L1 backgrounds in terms of self-representation.

These findings underscore the importance of emulating international journal practices to improve academic writing among Chinese EFL learners. Specifically, by adopting participant-oriented bundles and carefully analyzing subtle differences in usage patterns, learners can cultivate a stronger academic identity and produce research papers that communicate more persuasively. To substantiate these recommendations, future research should adopt a detailed, controlled intervention study to verify these recommendations. For example, one study might recruit 50–60 Chinese EFL learners from a single institution and randomly assign them to an experimental group and a control group. The experimental group would receive an eight - week intervention specifically designed to teach academic writing strategies drawn from international journal practices. This instruction might include:

Weekly 45 - minute sessions focusing on explicit training in participant-oriented bundles, critical metadiscourse features (e.g., hedges, boosters, and self-mention markers), and analysis of authentic academic texts.

Practical exercises such as annotating exemplar journal articles, engaging in peer review with structured checklists, and drafting introduction sections of research papers using these strategies.

Supplementary activities such as reflective journaling or focus groups to capture changes in students' self-perceived academic identity and voice.

Both groups would complete pre-test and post-test writing tasks that are evaluated using a standardized rubric, along with automated and manual frequency counts of targeted rhetorical markers. Quantitative data—using paired t-tests or ANCOVAs to control for initial differences—would compare writing quality, effective metadiscourse deployment, and self-representation across the two groups. Additionally, qualitative data from interviews or open-ended surveys could provide insights into learners' perceptions and the practical challenges they face in adopting these strategies.

This specific design would yield robust empirical evidence regarding whether explicit, targeted instruction leads to measurable improvements in academic writing, thereby substantiating the pedagogical recommendations for enhancing EAP instruction among Chinese EFL learners.



## CONCLUSIONS

Our investigation into metadiscourse in academic writing has yielded a multifaceted understanding of how Chinese EFL learners and international journal experts negotiate interpersonal meaning and textual structure. Drawing on extensive corpus analyses (CLMA\_C and ILJA\_C) and building on prior theoretical frameworks (e.g., Hyland, 2005), our study conceptualizes metadiscourse as a continuum. At one end, metadiscourse is narrowly construed as devices for textual organization and reflexivity (e.g., signaling structural elements), while at the other, it is seen as a comprehensive set of interpersonal strategies that articulate the writer's stance, engage readers, and construct academic identity. This continuum demonstrates that while metadiscourse lacks uniform nomenclature and sharply defined boundaries, its resources vary in their degree of manifestation. The more limited view, which emphasizes structural referentiality, contrasts with broader interpretations that encompass self-reference, reader engagement, and the nuanced interplay between authorial voice and audience expectations.

Our corpus-based analysis further refines this model by revealing significant quantitative and qualitative differences between Chinese EFL learners and international journal experts. Specifically, although learners show a diverse repertoire of certainty stance adverbs, maximizers, hedging devices, self-mentions, and engagement markers, their frequency of use and collocational patterns differ markedly from those observed in expert texts. For instance, learners use active-voice constructions and favor placing adverbs like 'actually' at the start of sentences to emphasize factual certainty; in contrast, experts tend to adopt passive-voice structures and pair adverbs such as 'clearly' with content-specific nouns (e.g., 'the figure', 'the table') that align with established academic conventions. With the maximizer 'fully', learners display a tendency toward collocational patterns and semantic prosody influenced by native language

transfer and interlanguage development. Whereas experts employ ‘fully’ to denote comprehensive engagement with theories or data in a commendatory tone, learners more often use it to signal personal desire for complete understanding. In the realm of hedging, despite similar diversity in speculative verbs between the two groups, our findings reveal that learners rely more heavily on impersonal constructions and exhibit notable discrepancies in the frequency and selection of quotative and sensorial verbs. For example, learners favor verbs like ‘claim’ over ‘argue’ and demonstrate restricted use of complex collocational structures (such as ‘it would seem to be/that-clause’) that experts routinely deploy. Similarly, our analysis of self-mention strategies shows that while both groups use the pronoun ‘we’ to guide discourse, experts integrate it with content-rich modifiers to assert evaluative positions and innovations, whereas learners more frequently attach generic function words that serve primarily to maintain logical cohesion.

Theoretically, our study advances existing models of metadiscourse by incorporating cross-cultural and interlanguage dimensions into the continuum framework. Whereas previous research (e.g., Hyland, 2005) has focused predominantly on categorizing metadiscourse in terms of structural versus interpersonal functions, our work demonstrates how cultural and academic norms influence not only the frequency but also the collocational deployment of these markers. By systematically comparing two distinct corpora, we show that the mastery of metadiscourse features is not merely a matter of linguistic competence but also reflects broader cultural conventions and academic practices. This nuanced perspective expands the theoretical models of metadiscourse by arguing that the continuum is dynamically modulated by factors such as native language influence, academic socialization, and genre-specific expectations. In doing so, our findings underscore the need for a more integrative model that accounts for variability in both production and usage across different learner groups and disciplinary contexts.

Besides, this dissertation analyzes the interplay between metadiscourse and these modes of persuasion, focusing on three strategies: rhetorical hyping (amplifying claims to emphasize significance), hedging (mitigating assertions to express caution or uncertainty), and authorial identity construction (crafting a credible persona to bolster trust). These strategies exemplify how metadiscourse operates as a dynamic mechanism for persuasion, balancing logical rigor, ethical credibility, and emotional resonance.

Pedagogically, our research has clear implications for EAP instruction. The observed discrepancies between Chinese EFL learners and expert writers suggest that targeted, corpus-based teaching approaches can play a crucial role in bridging the gap. For instance, our findings indicate that explicit training on the strategic use of participant-oriented chunks—such as employing engagement markers and self-mentions that align with disciplinary conventions—can help learners construct a more persuasive academic voice. In practice, EAP instructors might incorporate activities that use authentic corpora to highlight how experts employ metadiscourse elements. Specific exercises could include:

**Annotation tasks:** Learners analyze and annotate journal articles to identify instances of certainty stance adverbs, maximizers, hedging verbs, and self-mentions, thereby internalizing expert patterns.

**Comparative analysis:** By comparing excerpts from their own writing with those drawn from expert corpora, students can pinpoint discrepancies in collocational patterns (e.g., the use of ‘clearly’ with specific nouns versus generic pronouns) and adjust their strategies accordingly.

**Guided revision workshops:** Instructors can design revision sessions where learners receive corpus-informed feedback on their metadiscourse usage, focusing on adopting structures typical of expert writing, such as shifting from active to passive voice where appropriate or employing complex hedging constructions.

Reflective journaling: Students document changes in their understanding of metadiscourse functions, fostering awareness of how these elements contribute to authorial identity and textual coherence.

These corpus-based strategies not only ground theoretical insights in practical application but also provide a replicable framework for enhancing academic writing skills among EFL learners.

While our study offers robust evidence of metadiscourse differences between learners and experts, several avenues for future research remain. We recommend that subsequent studies consider:

Exploring spoken metadiscourse: Given that academic discourse increasingly involves multimodal communication (e.g., presentations, seminars), future investigations could examine how metadiscourse markers are deployed in spoken academic contexts. This would extend our findings from written texts to dynamic, real-time interactions, providing insights into the interactivity and immediacy of academic communication.

Expanding to additional L1 backgrounds: To determine the cross-cultural generalizability of our results, researchers should replicate our study with EFL learners from different first-language backgrounds (e.g., Arabic, Spanish, or European languages). Such comparative analyses would help to clarify the extent to which native language influences metadiscourse usage and inform tailored pedagogical interventions.

Longitudinal and intervention studies: Future research might adopt longitudinal designs to track changes in metadiscourse usage over time, particularly as learners progress from novice to expert-like writing. Additionally, intervention studies that implement targeted corpus-based instruction, similar to the pedagogical approaches recommended above, can empirically validate the efficacy of explicit metadiscourse training on writing outcomes.

Multi-modal discourse analysis: Integrating both written and spoken academic discourse analyses will provide a holistic view of how metadiscourse

functions across different communicative modes. Such studies could investigate whether similar patterns of difference emerge in the oral presentations or classroom interactions of EFL learners compared to expert speakers.

### **Uniqueness and advancement of theoretical models.**

The uniqueness of our findings lies in the detailed corpus-based contrast between Chinese EFL learners and international experts—a comparison that illuminates not only quantitative frequency differences but also qualitative divergences in collocational patterns and syntactic constructions. Our work advances theoretical models of metadiscourse by demonstrating that:

The deployment of metadiscourse elements is not monolithic but varies systematically along a continuum shaped by cultural and academic norms.

Even when learners possess a broad lexical repertoire, the nuanced usage—reflected in choices between active versus passive constructions, or in the selection of specific hedging verbs—differentiates expert-like writing from novice performance.

The interplay between different types of metadiscourse markers (certainty stance adverbs, maximizers, hedging devices, etc.) creates a complex rhetorical fabric that underpins effective academic communication, suggesting that theoretical models must account for both individual marker functions and their synergistic effects in establishing authorial identity.

By integrating these dimensions into a cohesive theoretical framework, our study not only reinforces but also extends current conceptions of metadiscourse, providing a robust basis for both further academic inquiry and targeted pedagogical practice.

In conclusion, our research synthesizes existing theoretical perspectives on metadiscourse with new corpus-based evidence, revealing that Chinese EFL learners and international journal experts differ significantly in both the quantitative and qualitative aspects of metadiscourse usage. These differences are driven by cultural, linguistic, and academic factors, and they have practical

implications for improving academic writing through explicit, corpus-informed instruction in EAP contexts. By offering specific recommendations for future research—such as investigating spoken metadiscourse and including learners from a wider range of L1 backgrounds—we provide a clear roadmap for advancing the field. Ultimately, our study not only deepens theoretical understanding of metadiscourse but also offers actionable insights that can enhance academic writing pedagogy, thereby contributing to the broader domains of comparative linguistics, academic discourse analysis, and second language acquisition.

## REFERENCES

1. Abbuhl, R. (2006). Hedging and boosting in advanced-level L2 legal writing: The effect of instruction and feedback. In: *Educating for advanced foreign language capacities: Constructs, curriculum, instruction, assessment*. Washington, DC: Georgetown University Press. pp. 152-164.
2. Abdi, R., Rizi, M. T., & Tavakoli, M. (2010). The cooperative principle in discourse communities and genres: A framework for the use of metadiscourse. *Journal of pragmatics*, 42(6), 1669-1679.
3. Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. Amsterdam: John Benjamins Publishing Company. pp. 47-55.
4. Adrian, D., & Fajri, M. S. A. (2023). Hedging practices in soft science research articles: A corpus-based analysis of Indonesian authors. *Cogent Arts & Humanities*, 10(1).
5. Aguilar, M. (2008). *Metadiscourse in academic speech: A relevance-theoretic approach*. Lausanne: Peter Lang. pp. 205-237.
6. Alghazo, S., al Salem, M. N., & Alrashdan, I. (2021). Stance and engagement in English and Arabic research article abstracts. *System*, 103, 102681.
7. Almakrob, A. Y. (2020). Native versus Nonnative English writers' use of hedging in linguistics dissertations. *Asian EFL Journal*, 27(4.4), 360-381.
8. Alrajhi, M. (2019). The semantic prosody and semantic preference of maximizers in Saudi EFL writings. *International Journal of Language and Linguistics*, 6(3), 30-39.
9. Alonso, R. A., Alonso, M. A., & Mariñas, L. T. (2012). Hedging: An exploratory study of pragmatic transfer in nonnative English readers' rhetorical preferences. *Ibérica*, (23), 47-64.

10. Altenberg, B. (1991). Amplifier collocations in spoken English. In S. Johansson (ed.), *English Computer Corpora: Selected Papers and Research Guide*. Berlin: De Gruyter Mouton. pp.127-130.
11. Anthony, L. (2019). Tools and strategies for Data-Driven Learning (DDL) in the EAP writing classroom. In K. Hyland & L. Wong (Eds.), *Specialised English: New directions in ESP and EAP research and practice*. Routledge. pp. 179–194.
12. Aull, L. L., & Lancaster, Z. (2014). Linguistic markers of stance in early and advanced academic writing: A corpus-based comparison. *Written communication*, 31(2), 151-183.
13. Bahns, J. (1993). Lexical collocations: A contrastive view. *ELT journal*, 47(1), 56-63.
14. Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273–306.
15. Bakhtin, M. M. (1996). *From the archival records to the “Problem of speech genres”*. Moscow: Russian dictionaries. pp. 159-207. (In Russ.).
16. Bateson, G. (1972). *Steps to an ecology of mind*. New York: Ballantine. pp. 59-67.
17. Beauvais, P. J. (1989). A speech act theory of metadiscourse. *Written communication*, 6(1), 11-30.
18. Bennett, K. (2009). English academic style manuals: A survey. *Journal of English for academic purposes*, 8(1), 43-54.
19. Benwell, B. (2006). *Discourse and identity*. Edinburgh: Edinburgh University Press. pp. 17-19.
20. Biber, D., & Finegan, E. (1988). Adverbial stance types in English. *Discourse processes*, 11(1), 1-34.



21. Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education.
22. Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for academic purposes*, 5(2), 97-116.
23. Bloch, J. (1995). A comparison of the use of citations in Chinese and English academic discourse. *Academic writing in a second language*, 6(1), 231-274.
24. Blommaert, J. (2005). *Discourse: A critical introduction*. Cambridge University Press.
25. Boginskaya, O. A. (2022). Functional categories of hedges: A diachronic study of Russian research article abstracts. *Russian Journal of Linguistics*, 26(3), 645-667.
26. Brezina, V., Weill-Tessier, P. and McEnery, A. (2020). Lancsbox (6.0) [Software]. Retrieved from <http://corpora.lancs.ac.uk/lancsbox/>
27. Brown, G., and George Y. (1983). *Discourse analysis*. Cambridge: Cambridge University Press. pp. 27-67.
28. Bu, J. (2014). Towards a pragmatic analysis of metadiscourse in academic lectures: From relevance to adaptation. *Discourse Studies*, 16(4), 449-472.
29. Cabanes, P. P. (2007). A contrastive analysis of hedging in English and Spanish architecture project descriptions. *Revista española de lingüística aplicada*, (20), 139-158.
30. Canagarajah, S. (2002). *Critical academic writing and multilingual students*. University of Michigan Press.
31. Canagarajah, S. (2013). *Translingual practice: Global Englishes and cosmopolitan relations*. Routledge.
32. Caulfield, T. (2018). Spinning the genome: why science hype matters. *Perspectives in Biology and Medicine*, 61(4), 560-571.
33. Caulfield, T., & Condit, C. (2012). Science and the sources of hype. *Public Health Genomics*, 15(3-4), 209-217.

34. Charles, M. (2006). Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines. *English for Specific Purposes*, 25(3), 310-331.
35. Charles, M., Pecorari, D., & Hunston, S. (2009). Introduction: Exploring the interface between corpus linguistics and discourse analysis. *Academic writing. At the interface of corpus and discourse*, pp. 1-13.
36. Chen, C., & Zhang, L. J. (2017). An intercultural analysis of the use of hedging by Chinese and Anglophone academic English writers. *Applied Linguistics Review*, 8(1), 1-34.
37. Conley, T. M. (1984). The enthymeme in perspective. *Quarterly Journal of Speech*, 70(2), 168-187.
38. Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second-language writing*. Cambridge University Press.
39. Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for specific purposes*, 23(4), 397-423.
40. Coxhead, A.J. (1998). *The development and evaluation of an academic word list*. Unpublished M.A. Thesis. Wellington: Victoria University of Wellington.
41. Coxhead, A.J., & Nation, P. (2001). The specialised vocabulary of English for academic purposes. *Research perspectives on English for academic purposes*, 252-267.
42. Crismore, A. (1989). *Talking with readers: Metadiscourse as rhetorical act*. New York: Peter Lang. pp.282.
43. 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.

44. Crismore, A., & Farnsworth, R. (1989). Mr. Darwin and his readers: Exploring interpersonal metadiscourse as a dimension of ethos. *Rhetoric Review*, 8(1), 91-112.
45. Crismore, A., & Farnsworth, R. (1990). Metadiscourse in popular and professional science discourse. In *The writing scholar*. Newbury Park: Sage Publications. pp.118-136.
46. Curry, M. J., & Lillis, T. (Eds.). (2017). *Global academic publishing: Policies, perspectives and pedagogies* (Vol. 1). Multilingual Matters.
47. Dafouz-Milne, E. (2008). The pragmatic role of textual and interpersonal metadiscourse markers in the construction and attainment of persuasion: A cross-linguistic study of newspaper discourse. *Journal of Pragmatics*, 40(1), 95-113.
48. DeCarrico, J., & Nattinger, J. R. (1988). Lexical phrases for the comprehension of academic lectures. *English for specific purposes*, 7(2), 91-102.
49. Dudley-Evans, T. (2002). Genre analysis: An approach to text analysis for ESP. In *Advances in written text analysis*. London: Routledge pp. 233-242.
50. Dugalich, N. M., Han, H. (2024). Certainty stance adverbs in Chinese linguistic academic writing: A corpus-based study. *RUDN Journal of Language Studies, Semiotics and Semantics*, 15(1), 248-261.
51. Fairclough, N. (2013). Critical discourse analysis. In *The Routledge handbook of discourse analysis*. Routledge. pp. 9-20.
52. Ecklund, E. H., Johnson, D. R., & Matthews, K. R. (2015). Commentary: Study highlights ethical ambiguity in physics. *Physics Today*, 68(6), 8-10.
53. Fitriati, S. W., & Gayatri, N. M. (2021). Thematic Progression in EFL Learners' Writing: A Literature Review. *Lingua Cultura*, 15(2), 257-262.
54. Flower, L. (1987). *Problem-solving strategies for writing* (3rd ed.). New York: Harcourt. pp. 294.
55. Flowerdew, L. (2004). The problem-solution pattern in apprentice vs. professional technical writing: an application of appraisal theory. *Corpora and language learners*. Amsterdam: John Benjamins. pp.125-135.

56. Flowerdew, J. (2008). Scholarly writers who use English as an additional language: What can Goffman's "Stigma" tell us? *Journal of English for Academic Purposes*, 7(2), 77–86.
57. Flowerdew, L. (2015). Corpus-based research and pedagogy in EAP: From lexis to genre. *Language Teaching*, 48(1), 99-116.
58. Fløttum, K. (2007). *Language and discipline perspectives on academic discourse*. Newcastle upon Tyne, UK: Cambridge Scholars Publishing. pp.280.
59. Fraser, B. (2010). Pragmatic competence: The case of hedging. In G. Kaltenböck, W. Mihatsch, & S. Schneider (Eds.), *New approaches to hedging*. Bingley: Emerald. pp. 15-33.
60. Gilbert, G. N., & Mulkay, M. (1984). *Opening Pandora's box: A sociological analysis of scientists' discourse*. Cambridge: Cambridge University Press. pp. 39.
61. Goffman, E. (1974). *Frame analysis*. New York: Harper & Row. pp.1-40.
62. Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and lexical phrases. *Phraseology: Theory, analysis and applications*, Oxford: Clarendon Press. pp. 145-160.
63. Grice, H.P. (1975). *Logic and conversation*. In P. Cole, & J. L. Morgan (Eds.), *Speech acts*. New York: Academic. pp. 26-40.
64. Gu, X., Xu, Z. (2021). Sustainable Development of EFL Learners' Research Writing Competence and Their Identity Construction: Chinese Novice Writer-Researchers' Metadiscourse Use in English Research Articles. *Sustainability*, 13(17), 9523.
65. Halliday, M. A. K. (1973). *Explorations in the functions of language*. NY: Elsevier-North Holland. pp.143.
66. Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Edward Arnold. pp.58-82.
67. Halliday, M. A. K., & Matthiessen, C. M. I. M. . (2014). *Halliday's introduction to functional grammar*. Routledge.

68. Harris, Z.S. 1959. Linguistics transformations for information retrieval. In *papers in Structural and Transformational Linguistics*. Dordrecht: D. Reidel, pp.464-466.
69. Haufiku, N. K. T., & Kangira, J. (2018). An exploration of hedging and boosting devices used in academic discourse focusing on English theses at the University of Namibia. *Studies in English Language Teaching*, 6(1), 1-11.
70. Hinkel, E. (2004). Tense, aspect and the passive voice in L1 and L2 academic texts. *Language teaching research*, 8(1), 5-29.
71. Hoey, M. (2001). *Textual interaction: An introduction to written discourse analysis*. Portland: Psychology Press. pp.35-50.
72. 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.
73. Hu, G., & Cao, F. (2015). Disciplinary and paradigmatic influences on interactional metadiscourse in research articles. *English for Specific Purposes*, 39, 12-25.
74. Hunston, S. (1993). Professional conflict–Disagreement in academic discourse. In *Text and technology*. John Benjamins. pp. 115
75. Hunston, S. (1994). Evaluation and organization in a sample of written academic discourse. In: Coulthard, M. (ed.) *Advances in written text analysis*. London: Routledge. pp. 205-232.
76. Hunston, S. (1995). A corpus study of some English verbs of attribution. *Functions of language*, 2(2), 133-158.
77. Hyland, K. (1996). Nurturing hedges in the ESP curriculum. *System*, 24(4), 477-490.
78. Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam: John Benjamins. pp. 126.
79. Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of pragmatics*, 30(4), 437-455.

80. Hyland, K. (2001). Bringing in the reader: Addressee features in academic articles. *Written communication*, 18(4), 549-574.
81. Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for specific purposes*, 20(3), 207-226.
82. Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of pragmatics*, 34(8), 1091-1112.
83. Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of second language writing*, 12(1), 17-29.
84. Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor: University of Michigan Press. pp. 41-61.
85. Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. United Kingdom: Bloomsbury Publishing. pp. 48-111
86. Hyland, K. (2017). Learning to write for academic purposes: Specificity and second language writing. In: *Teaching writing for academic purposes to multilingual students*. London: Routledge. pp. 24-41.
87. Hyland, K. (2017). Metadiscourse: What is it and where is it going? *Journal of Pragmatics*, 113, 16-29.
88. Hyland, K., & Jiang, F. K. (2016). “We must conclude that...”: A diachronic study of academic engagement. *Journal of English for Academic Purposes*, 24, 29-42.
89. Hyland, K., & Jiang, F. K. (2018). “In this paper we suggest”: Changing patterns of disciplinary metadiscourse. *English for specific purposes*, 51, 18-30.
90. Hyland, K., & Jiang, F. (2020). Academic discourse and global publishing: Disciplinary persuasion in changing times. *Applied Linguistics*, 41(1), 1-22.
91. Hyland, K., & Jiang, F. K. (2021). ‘Our striking results demonstrate...’: Persuasion and the growth of academic hype. *Journal of Pragmatics*, 182, 189-202.
92. Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students’ writing. *Journal of second language writing*, 6(2), 183-205.

93. Hyland, K., & Hamp-Lyons, L. (2002). EAP: Issues and directions. *Journal of English for academic purposes*, 1(1), 1-12.
94. Hyland, K., & Tse, P. (2004). "I would like to thank my supervisor". Acknowledgements in graduate dissertations. *International Journal of Applied Linguistics*, 14(2), 259-275.
95. Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, 25(2), 156-177.
96. Ifantidou, E. (2005). The semantics and pragmatics of metadiscourse. *Journal of pragmatics*, 37(9), 1325-1353.
97. Ivanič, R. (1998). *Writing and identity: The discursive construction of identity in academic writing*. John Benjamins.
98. Ivanič, R., & Camps, D. (2001). I am how I sound: Voice as self-representation in L2 writing. *Journal of second language writing*, 10(1-2), 3-33.
99. Jakobson, R. (1985). *Verbal art, verbal sign, verbal time*. Minneapolis: University of Minnesota Press. pp. 143-205.
100. Jiang, F. (2016). Stance construction and interpersonal interaction of shell nouns. *Contemporary Foreign Languages*, 39(04), 470-482..
101. Jiang, F. (2017). Stance and voice in academic writing: The "noun+ that" construction and disciplinary variation. *International Journal of Corpus Linguistics*, 22(1), 85-106.
102. Jiang, F. (2019). *Corpora and EAP studies*. Beijing: Foreign Language Teaching and Research Press. pp. 9-14.
103. Jiang, F. (2020). A diachronic multi-dimensional investigation into the stylistic features in academic discourse. *Foreign Language Teaching and Research*, 52 (5), 663-673+798.
104. Jiang, F. (2023). A diachronic disciplinary study of rhetorical hypes used in English and Chinese academic discourse. *Foreign Languages and Their Teaching*, 328(01), 24-25.

105. Jiang, F. K., & Hyland, K. (2020). Prescription and reality in advanced academic writing. *Ibérica*, (39), 14-42.
106. Jiang, F. K., & Ma, X. (2019). Positioning and proximity of reader engagement. *Engagement in Professional Genres*, 301, 29.
107. Jiang, F. K., & Wang J.J. (2021). Lexico-grammatical profile of self-mention used by Chinese students in research writing. *Foreign Languages in China*, 18(5), 90-97.
108. Johns, T. (1986). Micro-concord: A language learner's research tool. *System*, 14(2), 151-162.
109. Jordan, R. R. (1997). *English for Academic Purposes: A Guide and Resource Book for Teachers*. Cambridge: Cambridge University Press. pp. 1-6.
110. Karasik, V. I. (2020). Utterance, Genre, Discourse: Semiotic Modeling. *Speech Genres*, 2 (26), 90–99 (in Russian). <https://doi.org/10.18500/2311-0740-2020-2-26-90-99>
111. Kashiha, H. (2024). Stance-taking in peer reviewer and thesis examiner feedback on Iranian scholarly contributions. *Journal of English for Academic Purposes*, 68, 101364.
112. Keller, E. (1979). Gambits: Conversational strategy signals. *Journal of pragmatics*, 3(3-4), 219-238.
113. Kreutz, H., & Harres, A. (1997). Some observations on the distribution and function of hedging in German and English academic writing. *Trends in linguistics studies and monographs*, 104, 181-202.
114. Kumpf, E. P. (2000). Visual metadiscourse: Designing the considerate text. *Technical communication quarterly*, 9(4), 401-424.
115. Kopple, W. V. (1985). Some exploratory discourse on metadiscourse. *College Composition & Communication*, 36(1), 82-93.
116. Kopple, W. V., & Shoemaker, A. (1988). Metadiscourse and the recall of modality markers. *Visible Language*, 22(2), 233.



117. Lakoff, G. (1973). Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*, 2(4), 458-508.
118. Lafuente - Millán, E. (2014). Reader engagement across cultures, languages and contexts of publication in business research articles. *International Journal of Applied Linguistics*, 24(2), 201-223.
119. Lancaster, Z. (2016). Expressing stance in undergraduate writing: Discipline-specific and general qualities. *Journal of English for Academic Purposes*, 23, 16-30.
120. Lee, J. J., & Casal, J. E. (2014). Metadiscourse in results and discussion chapters: A cross-linguistic analysis of English and Spanish thesis writers in engineering. *System*, 46, 39-54.
121. Lee, J. J., & Deakin, L. (2016). Interactions in L1 and L2 undergraduate student writing: Interactional metadiscourse in successful and less-successful argumentative essays. *Journal of second language writing*, 33, 21-34.
122. Li, T., & Wharton, S. (2012). Metadiscourse repertoire of L1 Mandarin undergraduates writing in English: A cross-contextual, cross-disciplinary study. *Journal of English for Academic Purposes*, 11(4), 345–356.
123. Li, Y. (2018). *A comparative study of English and Chinese academic book reviews in applied linguistics: The approaches of genre and metadiscourse* (PhD dissertation). Shanghai: Shanghai International Studies University. pp. 8-34.
124. Liang, M. C. (2010). Chi-square Test and Log-likelihood Ratio Calculation Tool (Software). Beijing: China Foreign Language Education Research Center.
125. Liu G. B., Wang F. Y. & Bingham Z. (2021). The Local Grammar Patterns of Evaluation in Chinese and Western Scholars' Research Articles. *Foreign Languages and Their Teaching*, (2), 62-70+149.
126. Lorenz, G. (2014). Overstatement in advanced learners' writing: stylistic aspects of adjective intensification. In *Learner English on computer*. London: Routledge. pp. 53-66.

127. Lou, B. C. (2020). Exploration of Corpus Application in Teaching Academic English for Graduate Students. *Degree and Postgraduate Education*, (7), 51-56.
128. Lu, G. (2020). A Study on Stance Markers in Undergraduate Theses of International Students. *Fudan Sinology Series*, (1), 76-85.
129. Luukka, M. R. (1994). Metadiscourse in academic texts. In B. L. Gunnarsson, P. Linell, & B. Nordberg (Eds.), *Text and talk in Professional Context*. Uppsala: ASLA (The Swedish Association of Applied Linguistics). pp.77-88.
130. Macleod, M. R., Michie, S., Roberts, I., Dirnagl, U., Chalmers, I., Ioannidis, J. P., ... & Glasziou, P. (2014). Biomedical research: increasing value, reducing waste. *The Lancet*, 383(9912), 101-104.
131. Mao, L. R. (1993). I conclude not: Toward a pragmatic account of metadiscourse. *Rhetoric Review*, 11(2), 265-289.
132. Malyuga, E. N., & Tomalin, B. (2024). Euphemisms in South African English economic discourse: Socio-cultural aspects. *Russian Journal of Linguistics*, 28(3), 512–534. – <https://doi.org/10.22363/2687-0088-39076>
133. Master, Z., & Resnik, D. B. (2013). Hype and public trust in science. *Science and engineering ethics*, 19, 321-335.
134. Mauranen, A. (1993). Contrastive ESP rhetoric: Metatext in Finnish-English economics texts. *English for specific Purposes*, 12(1), 3-22.
135. Mauranen, A. (1993). Cultural differences in academic rhetoric: A textlinguistic study. *Peter Lang*.
136. McEnery, T., & Nazareth A. K. (2002). Epistemic modality in argumentative essays of second-language writers. *Academic Discourse*, 182-195.
137. Millar, N., Salager-Meyer, F., & Budgell, B. (2019). “It is important to reinforce the importance of...”: ‘Hype’ in reports of randomized controlled trials. *English for Specific Purposes*, 54, 139-151.

138. Mu, C. (2010). A contrastive analysis of metadiscourse in Chinese and English editorials. *Foreign Language Learning Theory and Practice*, 4(6), 31-34.
139. Mu, C., Zhang, L. J., Ehrich, J., & Hong, H. (2015). The use of metadiscourse for knowledge construction in Chinese and English research articles. *Journal of English for Academic Purposes*, 20, 135-148.
140. Nash, W. (2021). *An uncommon tongue: The uses and resources of English*. London: Routledge. pp. 234.
141. Oakes, M. (1998). *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press. pp.24-27.
142. Ochodo, E. A., de Haan, M. C., Reitsma, J. B., Hooft, L., Bossuyt, P. M., & Leeftang, M. M. (2013). Overinterpretation and misreporting of diagnostic accuracy studies: evidence of “spin”. *Radiology*, 267(2), 581-588.
143. Özbay, A. Ş., & Aydemir, T. (2017). The Use of Maximizers and Semantic Prosodic Awareness of Tertiary Level Turkish EFL Learners. *Journal of Education and Practice*, 8, 40-50.
144. Pan, F. (2012). Stance adverbs in journal articles by Chinese and native writers: A pragmatic view. *Journal of PLA University of Foreign Languages*, 35(5), 9-12.
145. Partington, A. (2001). Patterns and meanings: using corpora for English language research and teaching. *International journal of corpus linguistics*, 6(1), 155-158.
146. Peacock, M. (2015). Stance adverbials in research writing. *Ibérica*, (29), 35-62.
147. Peacock, M., & Flowerdew, J. (Eds.). (2001). *Research perspectives on English for academic purposes*. Cambridge: Cambridge University Press. pp. 11-12.
148. Pennycook, A. (2018). *Posthumanist applied linguistics*. Routledge.

149. Peredrienko, T., & Balandina, E. (2022). The Intensifier *очень* in Russian Academic Discourse. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje*, 48(1), 171-196.
150. Pérez-Llantada, C. (2010). The discourse functions of metadiscourse in published academic writing: Issues of culture and language. *Nordic Journal of English Studies*, 9(2), 41-68.
151. Poos, D., & Simpson, R. (2002). Cross-disciplinary comparisons of hedging. In R. Reppen, S. Fitzmaurice, & D. Biber (Eds.), *Using corpora to explore linguistic variation*. Amsterdam: John Benjamins. pp. 3-23.
152. Ragan, S. L., & Hopper, R. (1981). Alignment in the job interview. *Journal of Applied Communication Research*, (9), 85-103.
153. Rossiter Jr, C. M. (1974). Instruction in metacommunication. *Communication Studies*, 25(1), 36-42.
154. Ruan, Z. (2020). Metadiscourse use in 12 student essay writing: A longitudinal cross-contextual comparison. *Chinese Journal of Applied Linguistics*, 42(4), 466-487.
155. Ryabtseva, N.K. (1992). Mental Performatives in Scientific Discourse. *Issues of Linguistics*, (4), 12-28. (In Russ.)
156. Salager-Meyer, F. (1997). I think that perhaps you should: A study of hedges in written scientific discourse. In T. Miller (Ed.), *Functional approaches to written text: Classroom applications*. Washington: English Language Programs-United States Information Agency. pp. 105-118.
157. Scarcella, R. (2003). *Academic English: A conceptual framework*. Irvine: University of California Linguistic Minority Research Institute. pp. 8-9.
158. Schiffrin, D. (1980). Meta-Talk: Organizational and Evaluate Brackets in Discourse. *Sociological inquiry*, 50 (3-4), 199-236.
159. Schmid, H. J. (2000). *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. Germany: De Gruyter. pp. 10-20.

160. Scott, M. (2010). WordSmith Tools (Software). Retrieved June 1, 2010, from University of Oxford, <http://www.lexically.net/downloads/version5/HTML/index.html>
161. Shafqat, A., Rafique, Memon, A., Tafseer, & Khan, A. (2022). Do Pakistani English writers hedge more in linguistics research than native English writers? *Journal of Humanities, Social and Management Sciences (JHSMS)*, 3(1), 243-257.
162. Shvedova, H. U. (1980). Russian Grammar. Vol. 1. Phonetics. Phonology. Stress. Intonation. Word Formation. Morphology. Moscow: Nauka. pp. 34-67.
163. Sinclair, J. (2004). *Trust the text*. London: Routledge. pp. 9-24.
164. Smirnova, E., & Strinyuk, S. (2020). Hedges in Russian EAP writing: A corpus-based study of research papers in management. *Journal of English as a Lingua Franca*, 9(1), 81-101.
165. Snow, C. E., & Uccelli, P. (2009). *The challenge of academic language*. The Cambridge handbook of literacy. Cambridge: Cambridge University Press. pp.112.
166. Sun, L. (2020). A Study on the Use of Metadiscourse and Identity Construction Features in Chinese Master's Academic English Writing. *Journal of Xi'an International Studies University*, 28(4), 28-33.
167. Swales, J. M., & Feak, C. B. (2012). *Academic Writing for Graduate Students: Essential Tasks and Skills*. 3rd ed. Ann Arbor, MI: University of Michigan Press. pp. 331.
168. Swales, J.M., Swales, J. (1990). *Genre analysis: English in academic and research settings*. Kiribati: Cambridge University Press. pp. 260.
169. Takimoto, M. (2015). A corpus-based analysis of hedges and boosters in English academic articles. *Indonesian Journal of Applied Linguistics*, 5(1), 95-105.

170. Tang, R., & John, S. (1999). The 'I' in identity: Exploring writer identity in student academic writing through the first person pronoun. *English for specific purposes*, 18, S23-S39.
171. Tessuto, G. (2011). Legal problem question answer genre across jurisdictions and cultures. *English for Specific Purposes*, 30(4), 298-309.
172. Thabet, R. A. (2018). A Cross-Cultural Corpus Study of the Use of Hedging Markers and Dogmatism in Postgraduate Writing of Native and Non-native Speakers of English. *Studies in Computational Intelligence*, 740, 677-710.
173. Thompson, G., & Yiyun, Y. (1991). Evaluation in the reporting verbs used in academic papers. *Applied linguistics*, 12(4), 365-382.
174. Kopple, W. J. V. (1980). *Experimental evidence for functional sentence perspective* (Doctoral dissertation). The University of Chicago. pp. 45.
175. Kopple, W. J. V. (1985). Some exploratory discourse on metadiscourse. *College Composition & Communication*, 36(1), 82-93.
176. Kopple, W. J. V. (1997). Refining and applying views of metadiscourse. Paper presented at the 48th annual meeting of the conference on college composition and communication. Phoenix: AZ. pp. 78.
177. Van Dijk, T. A. (2011). *Discourse studies: A multidisciplinary introduction*. SAGE Publications.
178. Van Dijk, T. A. (2008). *Discourse and power*. Palgrave Macmillan.
179. Varttala, T. (2001). *Hedging in Scientifically Oriented Discourse: Exploring Variatio*. Tampere: University of Tampere press. pp.64.
180. Vass, H. (2017). Lexical verb hedging in legal discourse: The case of law journal articles and Supreme Court majority and dissenting opinions. *English for Specific Purposes*, 48, 17-31.
181. Vassileva, I. (2001). Commitment and detachment in English and Bulgarian academic writing. *English for specific purposes*, 20(1), 83-102.

182. Vinkers, C. H., Tijdink, J. K., & Otte, W. M. (2015). Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: retrospective analysis. *British Medical Journal*, 351: 64-67.
183. Volkova, Y. A., & Panchenko, N. N. (2024). Motivational speech: Specifics of the speech genre functioning. *Speech Genres*, 19-1(41), 47–55. <https://doi.org/10.18500/2311-0740-2024-19-1-41-47-55>
184. Walková, M. (2019). A three-dimensional model of personal self-mention in research papers. *English for Specific Purposes*, 53, 60-73.
185. Wierzbicka, A. (1978). Metatext in Text (Based on the Polish Language). In: Nikolaeva, T. M. *New in Foreign Linguistics*. Vol. 8. Moscow: Progress. pp. 403-411.
186. Wang, H., & Chen, G. (2007). An investigation into the developmental features of Chinese EFL learners' use of amplifier collocations: A corpus-based approach. *Journal of Foreign Languages*, 30(1), 52-58.
187. Wang, J.J., & Jiang, F. (2019). Stance construction in academic English writing by Chinese PhD students of science and engineering. *Foreign Language World*, (3), 23-31.
188. Wang, J., & Zeng, L. (2021). Disciplinary Recognized Self-Presence: Self-Mention Used With Hedges and Boosters in PhD Students' Research Writing. *Sage Open*, 11(2).
189. Wei, N. X. (2016). Revisiting Academic English: Theory, Paths, and Methods. *Modern Foreign Languages*, 39 (2), 267-277+293.
190. Wei, Y. Y., & Lei, L. (2011). The use of amplifiers in the doctoral dissertations of Chinese EFL learners. *Chinese Journal of Applied Linguistics*, 34 (1), 47-61.
191. Williams, J. M. (1981). *Style: Ten lessons in clarity and grace*. Glenview, IL: Scott, Foresman. pp. 200-211.

192. Wishnoff, J. R. (2000). Hedging your bets: L2 learners' acquisition of pragmatic devices in academic writing and computer-mediated discourse. *Second Language Studies*, (19), 119-148.
193. Wodak, R., & Meyer, M. (2016). *Methods of critical discourse studies* (3rd ed.). SAGE.
194. Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied linguistics*, 21(4), 463-489.
195. Xu, F. (2015). Exploring stance markers in second language academic writing. *Foreign Languages and Their Teaching*, (5), 1-7.
196. Xu J. J. (2019). *Corpus and Discourse Studies*. Beijing: Foreign Language Teaching and Research Press. pp. 59-64.
197. Yagız, O., & Demir, C. (2014). Hedging Strategies in Academic Discourse: A Comparative Analysis of Turkish Writers and Native Writers of English. *Procedia - Social and Behavioral Sciences*, 158, 260–268.
198. Yang, Y. W., Xu, M.R.Z., & Liu, Q. H. (2019). A Comparative Study of Evidentiality in English and Chinese Academic Papers. *Journal of Xi'an International Studies University*, (3), 48-52.
199. Yoon, H. J., & Abdi Tabari, M. (2023). Authorial voice in source-based and opinion-based argumentative writing: Patterns of voice across task types and proficiency levels. *Journal of English for Academic Purposes*, 62, 101228.
200. Yu, W.S., & Zhong, Z. L. (2017) Corpus-based Research on Extended Meanings of English and Chinese Maximizers. *Foreign Language Education*, 38(5), 32-37.
201. Zanina, E. (2017). Move structure of research article abstracts on Management: Contrastive study (the case of English and Russian). *Journal of Language and Education*, 3(2), 63-72.
202. Zhang, D., & Sheng, D. (2023). A Contrastive Analysis of Metadiscourse by Native and EFL Lecturers in Chinese University MOOCs. In *New Trends on*



*Metadiscourse: An Analysis of Online and Textual Genres*. Cham: Springer International Publishing. (pp. 129-160)

203. Zhang, X. H. (2010). A Corpus-driven Study of Extended Units of Meaning: Taking Maximizers as an Example. *Foreign Languages and Their Teaching*, (4), 25-30.

204. Zhao, X. L. (2009). Expression of stance in Chinese learner English: Stance adverbs. *Journal of PLA University of Foreign Languages*, 32(5), 54-59.

## **DICTIONARIES**

205. Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (2000). *Longman grammar of spoken and written English*. Harlow: Pearson Education Limited. pp. 990.

206. Quirk, R., Greenbaum, G., Leech, G. & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman. pp. 589-593.