

Article

Sustainable Development of EFL Learners' Research Writing Competence and Their Identity Construction: Chinese Novice Writer-Researchers' Metadiscourse Use in English Research Articles

Xiaole Gu *  and Ziwei Xu

School of International Studies, Harbin Institute of Technology, Harbin 150001, China; xuziwei0710@163.com

* Correspondence: wanglingqingxl@sina.com

Abstract: English for foreign language (EFL) novice writer-researchers are faced with an increasing pressure for international publication as a prerequisite for sustainable career development in academia. The use of metadiscourse, as a key indicator for their discourse competence, has been a subject of research for English for Academic Purposes (EAP) and/or English for Specific Purposes (ESP) scholars. This study investigates metadiscourse features of research articles' (RA) results and discussion (R&D) sections written by Chinese PhD students and their writer identities reflected through metadiscourse choice. A corpus was built, consisting of a subcorpus of R&D of unpublished research articles (RAs) written by Chinese PhD students (CNWs) and one of the same part-genre by English-speaking expert writers (EEWs). Metadiscourse used by the two groups were identified based on Hyland's interpersonal model of metadiscourse. Quantitative analyses on the frequency and variety of metadiscourse markers found a significant difference not only in interactional metadiscourse but also in some subcategories of interactive and interactional metadiscourse, indicating that CNWs attach more importance to organisation of ideas than to the persuasiveness of arguments. A questionnaire survey was conducted to explore the influence of the CNWs' perception of RA writing on their metadiscourse choice. It revealed that knowledge of generic conventions and metadiscourse functions, awareness of the writer-reader relationship, and confidence in language competence may influence metadiscourse choice. The paper concludes with the view that the CNWs generally view themselves as a recounter and reporter of their research, remaining conservative when presenting an authoritative voice and a confident identity as a knowledge creator.

Keywords: metadiscourse; novice EFL writers; writer identity; research writing; interactive; interactional



Citation: Gu, X.; Xu, Z. Sustainable Development of EFL Learners' Research Writing Competence and Their Identity Construction: Chinese Novice Writer-Researchers' Metadiscourse Use in English Research Articles. *Sustainability* **2021**, *13*, 9523. <https://doi.org/10.3390/su13179523>

Academic Editors: Lawrence Zhang and Vincent T. Greenier

Received: 26 July 2021

Accepted: 20 August 2021

Published: 24 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

English for foreign language (EFL) novice writer-researchers are faced with an increasing pressure for international publication as a prerequisite for a sustainable career development in industry and academia. They need to be more cautious than English-speaking expert writers when looking for appropriate linguistic resources and rhetorical strategies to unfold new knowledge through texts as a knowledge creator on the one hand and negotiate existing discourses as a disciplinary community member on the other [1]. Metadiscourse contributes a great deal in these two aspects [2], thus attracting increasing attention from scholars in English for Academic Purposes (EAP) and/or English for Specific Purposes (ESP) writing. Metadiscourse is defined by Hyland as "aspects of the text which explicitly refer to the organisation of the discourse or the writer's stance towards either its content or the reader" [3] (p. 109). It plays the roles of organising the text, evaluating findings, engaging readers, and manifesting research significance in ways "that are meaningful and appropriate to a particular disciplinary community" [4] (pp. 438–440).

Although there are numerous studies on metadiscourse features in research articles (RAs) written by L1 writers, including those on disciplinary variations [5,6], cross-cultural

variations [7,8], or metadiscourse features of a particular category [9,10], studies targeting at RAs written by L2 novice writers are scarce [11]. EFL writers have been found to be inadequate in deploying metadiscourse resources [12–14]. There are several reasons for this. First, the use of metadiscourse can be quite flexible, involving a wide range of linguistic devices and serving a variety of functions [2]. Second, metadiscourse use often varies across cultures, and unawareness of such cross-cultural differences may lead to ineffective communication with potential readers for scholarly publication [15]. In addition, metadiscourse is often neglected in research and writing courses, which tend to focus on the lexico-grammar of propositional ideas rather than rhetorical strategies on interpersonal aspects [8]. Therefore, research into the metadiscourse use of EFL writers is highly needed to provide EAP/ESP instructors and novice writers with evidence and new perspectives on metadiscourse features in this research genre.

Previous work investigating metadiscourse features in different RA sections have mainly been focused on the abstract [9] and introduction [11], yet results and discussion (R&D) remains an under-researched part-genre. This study chooses the combined sections of results and discussion (R&D) of an RA as the focus due to its critical role in establishing the significance of the study based on research findings [16,17]. R&D is a place where the writer demonstrates their ability to think critically about his/her own research and those of others, and to convince readers of the value of the study [18]. Metadiscourse use unconforming to the purpose of this part-genre of a particular discipline may jeopardize the stimulating and convincing strength of the argument.

One aspect of RA metadiscourse features in EFL writings that remains underexplored is the relationship with writer identity. RA production involves the writer's conscious or unconscious construction of a writer identity by assuming both the roles of researcher and discourse constructor [19]. Studies reveal that EFL writers encounter not only challenges in deploying linguistic and rhetorical resources [12,20], but also in developing a scholarly voice [1,21,22].

This study aims to explore novice writers' metadiscourse features and identity construction, particularly targeted at Chinese engineering PhD students. This group of writers are generally required to publish at least 2–3 RAs in English-medium SCI-indexed journals in order for graduation, yet they often suffer rejections from targeted journals due to a lack of competence in the academic genre. Although universities in China have recently begun to incorporate research writing instruction into the graduate EAP curriculum, such instruction is often provided by inexperienced instructors who are still in the process of learning the genre-based pedagogies [8]. They mainly focus on basic, conventional rules of research writing while often neglecting lexico-grammar features and the relationship between these features and rhetorical purposes. Thus, it will be of importance to investigate in what areas and to what extent metadiscourse used by Chinese EFL writers of RAs is inadequate and how their metadiscourse use reflects their writer identity so as to facilitate genre-based EAP/ESP writing instruction.

The present study aims to achieve the following objectives: (1) To explore, through a corpus-based approach, metadiscourse features of the R&D of unpublished RAs written by Chinese PhD engineering students by comparing them with those displayed in the same part-genre by English expert writers in engineering; (2) To explore, through a questionnaire survey, whether Chinese novice writers' choice of metadiscourse is a result of their perception of RA writing; (3) How metadiscourse choice by these novice writer-researchers reflects the construction of their writer identity. Thus, the research questions are:

- (1) What are the metadiscourse features of the R&D of RAs written by Chinese novice writer-researchers?
- (2) Is Chinese novice writers' choice of metadiscourse a result of their perception of RA writing?
- (3) How does metadiscourse choice by these novice writer-researchers reflect the construction of their writer identity?

By combining both corpus-based approach and qualitative inquiry, this study hopes to offer inspirational findings on the relationship between metadiscourse use and writer identity construction.

2. Literature Review

2.1. Metadiscourse in RAs

The role of metadiscourse in “facilitating communication, supporting a writer’s position and building a relationship with an audience” has been increasingly recognised in research writing [4] (p. 438). Successful research writers utilize a wide range of metadiscourse resources to achieve informational clarity on the one hand and guide readers to desired interpretations and a manifest stance on the other [23]. By doing so, they also strive to achieve “a balance between objective information, subjective evaluation and interpersonal negotiation” [24] (p. 294) so as to meet discourse community expectations and to gain acceptance for their statements [7].

Studies have tended to adopt a corpus-based approach to explore metadiscourse features in RAs, normally by comparing RAs with other genres, across disciplines, and across cultures. Swales [25] and Bunton [26] found that RAs engage in less metadiscourse than dissertation writing. More specifically, Kawase, by comparing various types of metadiscourse in the introduction section of RAs and PhD dissertations, revealed that RAs tended to be less explicit in their exposition than dissertations, and speculated that variations in metadiscourse use between these two part-genres may be attributed to different writer–reader relationships [11].

It is generally agreed that metadiscourse in RAs is “socially authorised and contextually constrained by the disciplinary communities in which it occurs” [4] (p. 448). Numerous studies have found differences across disciplines in the use of interactive metadiscourse (e.g., [5,27]) and interactional metadiscourse (e.g., [28,29]). Cao and Hu [5] and Hu and Cao [6] found differences in the use of several interactive and interactional metadiscourse resources between the natural and human sciences and attributed such differences to a different knowledge-knower structure between natural and human sciences. Hyland observed low frequencies of interpersonal metadiscourse in hard disciplines than in soft ones and speculated that there was a reluctance in the hard disciplines “to project a prominent authorial presence in presenting claims to their target community” [4] (p. 448).

Cross-cultural and cross-linguistic differences in metadiscourse use between RA writing in English and other languages are also manifested, since people from different cultures may assign different roles to writers, readers, and the text, thus using metadiscourse in different ways [2,7,23,30]. It is also revealed that, although metadiscourse elements are similar in texts written in different languages, strategies used to realize interpersonal functions seem to be “partly influenced and constrained by the lingua-cultural contexts in which texts are produced and consumed” [23] (p. 50). Comparative studies between English RAs and Chinese RAs have revealed that the former tend to use more interactional metadiscourse features to involve the audience in the text [7,31]. Writers from English-speaking cultures tend to adopt a more active role in making their ideas organized and explicit for readers to understand, while Chinese writers are found to be more implicit in expressing stance and engaging readers, passing the responsibility for interpreting the text to readers [8]. Li and Xu revealed that, compared with English RA writers, Chinese writers preferred metadiscourse markers for commenting on text to those for interacting with their readers, which contributed to “a more objective and detached style” in Chinese RAs than in English ones [8] (p. 54).

2.2. Metadiscourse in EFL Academic Writing

EFL novice writer-researchers’ metadiscourse use is a key indicator for their discourse competence and has been a subject of research for EAP/ESP writing scholars. Some studies have focused on the struggles that EFL writers encounter in their process of RA writing [22,32], while others concern the ability of these writers to convey a clear position of the

“author” on their arguments and to engage the reader directly in the texts [33,34]. Differences have been revealed in frequency and purpose of metadiscourse resources used by EFL learners from those by L1 English writers [30,35]. For instance, Lee and Deakin found that EFL academic essays generally contained fewer instances of interactional metadiscourse markers than L1 English ones and thus revealed less effective persuasion [35]. Hyland found that EFL writers had problems with both the type and range of metadiscourse markers, affecting the persuasive effect of their argument [4]. Specifically, EFL student writers were less strategic in stance-taking for establishing persuasive argumentation [36]. Zhang and Zhan suggested that the EFL learners’ voice development may be influenced by dynamic changes in culture, language, and education in a globalized society [37].

Metadiscourse use is found to affect the effectiveness of writing. RAs written by high-rated L2 writers often embrace more interactional metadiscourse markers, especially expressions such as hedges, attitudinal markers, and engagement markers, since, with such expressions, authors can establish good relationships with their potential readers [38,39]. Writing of successful EFL writers may display less of a reliance on interactive resources, a balanced use of interactional resources, and an increased range of metadiscourse markers, indicating a greater dialogic sense of interaction and audience involvement in advanced EFL learners’ writing [40,41]. They also use a higher variety of metadiscourse resources than less successful ones, who tend to be over-reliant on highly-frequent text connectives [41,42].

Despite a proliferation of studies on the metadiscourse features of English writing by Chinese EFL learners, most of them have focused on academic writing by undergraduates. Given the cross-genre differences in metadiscourse use, such studies may not provide direct evidence or implications for novice writer researchers. Studies on RAs written by Chinese writers are mainly on published Chinese RAs in soft disciplines (e.g., applied linguistics [7], sociology [8]), while those on English RAs written by Chinese novice engineering researchers are rare. Considering that engineering researchers in China generally receive less formal training in research writing and are often under more pressure of international publication than those in soft disciplines, the writing of such a group of writers indeed calls for more attention. In addition, most findings on the influencing factors of metadiscourse use are mainly postulations based on quantitative findings, which lack support from writers’ reflection. It would be more inspirational if a qualitative inquiry was taken to explore how novice writers perceive RA writing and how such perceptions influence their decision-making about metadiscourse use.

2.3. Academic Writer Identity

According to Ivanič, “writing is an act of identity in which people align themselves with socio-culturally shaped possibilities of self-hood” [19] (p. 32). The social constructivist view sees writing as a social engagement in which writers interact with their potential readers, not only to convey messages, but to facilitate understanding [4]. In this process, writers also construct their own personality through their presentation of self and the way they explain their actions [28]. This means that writers adopt an appropriate identity through their choice of words to present ideas in ways that make sense to their readers and are acceptable by the disciplinary communities they belong to [43].

Despite the existence of disciplinary conventions, it is provincial to regard writer identity as fixed. Instead, writer identity is constantly changing due to the reality of heteroglossia, reflected by the dynamic, dialogic interactions between writers, texts, and the prospective audience [36]. Writer identity is not mono-faceted either, for different aspects of identity are constructed simultaneously in a text through lexico-grammatical choice [31]. Ivanič establishes a framework of writer identity, which contains four interrelated aspects: Autobiographical self (i.e., what a writer brings to the discourse from his past experience), discursual self (i.e., a writer’s self-representation constructed through discourse characteristics), authorial self (i.e., a writer’s presence or authoritativeness displayed in the text), and possibilities for self (i.e., possible identities available in the socio-cultural context of writing) [19]. Among these aspects, discursual self and authorial self are closely related

to discourse characteristics, such as metadiscourse, and are thus the focus of this study. In RAs, discursal identities are “constructed through the discourse characteristics of a text that reflect values, beliefs and power relations in the social context in which they were written” [19] (p. 25) and are associated with how writers are inclined to play various roles by choosing certain linguistic features. Authorial identity concerns how writers present self and establish authority for their claims in their writing in the sense of their position, opinions, and beliefs [19] (p. 27).

In RA writing, the presentation of discursal and authorial self is reflected in several roles that the writer takes by taking into consideration the contextual factors, including the nature of his research, discipline conventions, his power relationship with readers, and his sense of self [31]. The roles of a writer-researcher can be classified into six types: Conveyor of general knowledge, guide or navigator, conductor of research, evaluator of previous claims, originator of claims, and reflexive self [19,44]. These roles display an ascending order of authority, with knowledge conveyor being the least authoritative and reflexive self being the most.

A number of studies on EFL writing have investigated writers’ identity construction process by exploring specific metadiscourse markers. For instance, studies on authorial identity have revealed that English L1 speakers are more ready to present themselves explicitly in their RAs through relatively frequent use of self-mentions or stance and engagement markers, while writers of other languages display a covert author presence by avoiding the use of these metadiscourse markers, especially self-mentions [44,45]. Lee and Deakin, by exploring interactional metadiscourse in argumentative essays, found that Chinese EFL students were unwilling to construct an authoritative writer identity in their writing [35]. Similarly, Wu found that Chinese RA writers used stance and engagement markers less frequently than English RA writers, suggesting that the former were less willing to present a voice or community-recognized personality [31]. On the contrary, Geng and Warton, by investigating engagement resources, refuted the view that Chinese students were reluctant to critique in order to preserve and maintain public image [12]. These inconsistent findings point to the necessity to further explore the relationship between metadiscourse choice and identity construction, so as to provide more empirical evidence in this aspect.

3. Methods

3.1. Framework of Metadiscourse Markers

An interpersonal model of metadiscourse by Hyland [2] was adopted for data analysis in this study, since this model has been proved reliable for studying metadiscourse in research genres by many recent studies [6,7,35] and reflects the latest development in the methodology of metadiscourse analysis. The model classifies metadiscourse into two macro-types: Interactive metadiscourse (Table 1) and interpersonal metadiscourse (Table 2). Interactive metadiscourse refers to resources that “manage the information flow to explicitly establish his or her preferred interpretations” [3] (p. 138). It contains Transitions, Frame Markers, Endophoric Markers, Evidentials, and Code Glosses. The interactional metadiscourse concerns “the writer’s efforts to control the level of personality in a text and establish a suitable relationship to his/her data, arguments, and audience, marking the degree of intimacy, the expression of attitude, the communication of commitments, and the extent of reader involvement” [3] (p. 141). It consists of Hedges, Boosters, Attitude Markers, Self-mentions, and Engagement Markers.

Table 1. Classification of interactive metadiscourse.

Categories	Function	Subcategories
Transitions	express internal relationships between discourse units	Addition (<i>in addition</i>), Contrast (<i>although</i>), Consequence (<i>therefore</i>)
Frame Markers	sequence text stages, announce discourse goals, and signal topic shifts	Sequencer (<i>firstly</i>), Label stages (<i>at this stage</i>), Announcer (<i>the purpose of this paper is</i>), Topicalizer (<i>now</i>)
Endophoric Markers	refer to previous or subsequent text	Non-linear (<i>see Figure 1</i>), Linear (<i>above, in Section 2.1</i>)
Evidentials	refer to the source of textual information cited from other texts	Integral citations (<i>X finds, according to X</i>), Non-integral (<i>(Smith, 2020)</i>)
Code Glosses	elaborate or restate propositional information	Exemplification (<i>for example</i>), Reformulation (<i>i.e.</i>)

Table 2. Classification of interactional metadiscourse markers.

Categories	Function	Subcategories
Hedges	withhold complete commitment to a proposition	Probability (<i>probably</i>), Inference (<i>suggest</i>), Approximate quantity (<i>approximately</i>), Approximate frequency (<i>sometimes</i>), Approximate degree (<i>somewhat</i>), Approximate limitation (<i>mainly</i>)
Boosters	close down alternatives and express certainty	Emphatic (<i>certainly</i>), Amplifying (<i>always</i>)
Attitude Markers	express personal views, attitudes, and evaluations	Obligation modals (<i>should</i>), Affective adverbs (<i>surprisingly</i>), Evaluative adjectives (<i>important</i>), Attitude verbs (<i>expect</i>), Attitude nouns (<i>advantage</i>)
Self-mentions	signal explicit author presence in the text	First-person singular (<i>I, me</i>), First-person plural (<i>we, us</i>), Third-person (<i>author</i>)
Engagement Markers	explicitly engage readers to get them involved in discourse	Directives (<i>notice</i>), Knowledge appeal (<i>as we all know</i>), Reader-mentions (<i>you</i>), Questions, Personal asides (<i>parenthesis</i>)

Each category is further classified in terms of more nuanced functions, following [5,6,23]. However, minor modifications were made in our study when classifying and identifying metadiscourse markers. First, instead of classifying Hedges into writer- and reader-oriented types as Lee and Casal [23], we classified them into subcategories of Probability, Inference, Approximate quantity, Approximate frequency, Approximate degree, and Approximate limitation, because this classification more clearly reflects salient functions of Hedging expressions. Second, the criteria proposed by Hyland [2] for identifying various types of metadiscourse markers were slightly altered. For example, Attitude Markers in our study included not only expressions concerning the writer's attitude towards his/her proposition (Example 1), but also those appraising the writer's or others' studies, which demonstrates the writer's effort to assert the value of his/her study (Example 2).

Example 1. *Interestingly, a similar type of carbon fibers . . . has been reported recently. (EEW10).*

Example 2. *In this study, all Cu based samples showed better field emission results, . . . , compared with their corresponding CuO-based samples. (EEW49).*

3.2. Corpora Construction

To study the metadiscourse features of Chinese novice writers (CNWs), we compared the results and discussion (R&D) of unpublished English RAs by Chinese PhD students and the same section of published RAs by English expert writers (EEWs). The unpublished RAs were collected from the writing center of a top Chinese university during 2018–2020. They were written by Chinese engineering PhD students of this university, who have all passed the English proficiency test for PhD students of the university and can be considered as intermediate–advanced EFL learners. They submitted their RA drafts to the writing center for revision with a hope of reaching the standard of their targeted international journals. The selection of RAs for the present study observes three criteria: (1) Each selected RA should have the conventional IMRD structure (though subtitles for each section may be different), with an integrated R&D section; (2) Each selected RA should be linguistically comprehensible, with grammatical errors and inappropriate uses not severely interfering with the reader’s understanding of the RA; (3) Student-writer of the selected RA has received little formal training in RA writing and has little experience in publishing RAs (revealed in their application form for RA revision). Fifty RAs were selected and produced a CNWs’ R&D corpus of 73,165 words. For comparison, a parallel EEWs’ R&D corpus of published RAs was built, containing 92,584 words. These RAs were selected by the Chinese PhD students from international journals with high impact factors (JCR Q1 or Q2) to which they intended to submit their RAs. They were similar to the selected unpublished RAs in discipline, research topic, and generic structure, and were written by renowned scholars who are either L1 English speakers or affiliated within English-speaking institutions. A description of the two subcorpora is given in Table 3.

Table 3. Description of the two subcorpora.

	Unpublished RAs by CNWs	Published RAs by EEWs
No. of RAs	50	50
Total length of RAs	305,829	387,001
Total tokens in R&D	73,165	86,768
Average length of R&D	1463.3	1851.68
Range of the length of R&D	335–4758	407–4672

3.3. Analysis of Corpora

A word or a chunk of words expressing a metadiscursive function serves as the unit of analysis. Expressions with similar patterns, such as *Table 1* and *Figure 1*, were counted as the same metadiscourse marker. By referring to the classification in Section 3.1, all the metadiscourse markers were manually identified and annotated, since metadiscourse is highly contextual and a particular expression can serve different functions. For example, the modal verb *would* serves metadiscursive function when it refers to the result or effect of a possible situation, but is not counted as a metadiscourse marker when it is used as the past form of *will* to project a future happening in the past. Another example is the conjunction *and*, which serves a metadiscursive function when connecting two clauses to indicate a clausal relationship, but not when connecting two words or phrases.

The two authors first identified metadiscourse markers of each subcategory in an R&D sample together against the definitions given by Hyland [2] (pp. 48–54) to make sure they understood and agreed on the coding criteria. Any inconsistencies during the coding process were discussed. They then annotated the data individually. Intercoder reliability was measured and a coding reliability coefficient of 0.82 was obtained, indicating a relatively high level of agreement between coders on the categorization of metadiscourse markers.

Quantitative analyses were conducted of the coded metadiscourse markers. The annotated text files were uploaded to Antconc (v. 3.2.4) for metadiscourse frequency count. Identified metadiscourse markers of each subcategory were counted to reveal the writers’ metadiscourse repertoire and preference. They were then normalized to 10,000 words for comparison between the subcorpora. Chi-square analyses, a non-parametric test commonly

used in corpus research, were undertaken to determine any statistically significant differences in the frequency of metadiscourse categories between the two subcorpora. p value was set at 0.05 for all statistical tests.

3.4. Questionnaire Survey

An online questionnaire was also conducted among the CNWs, aiming to explore their perception of metadiscourse in RAs' R&D. The questionnaire focused on perceptions of RAs' R&D functions, writer-reader interactions, and metadiscourse use. The questionnaire contained two open-ended questions and 10 Likert-scale questions. The first open-ended question was on their views of the functions of R&D in RAs. The second was whether they would consider the potential readers when organising the information of their RAs and in what aspects they would adjust their RA's R&D in anticipation of readers' needs. The 10 Likert-scale questions were included to measure how often the CNWs think each category of metadiscourse should be used in RAs' R&D. A brief definition and a few typical examples of each category were given to make sure each respondent understood the terms. Each question contained 6 options: 0 stood for "I have no idea whether to use it", 1 for never, 2 for rarely, 3 for occasionally, 4 for quite often, and 5 for always. The respondents were also asked to provide their reasons for their choice. The questionnaire link was sent to the CNWs after their RAs were selected into the subcorpus. They were asked to do the survey before a deadline. Eventually, we collected responses from 47 students.

4. Results

4.1. Findings of the Corpus Study

To answer the first research question, i.e., what are the metadiscourse features of the R&D of RAs written by Chinese novice writer-researchers, we compared the frequency and variety of metadiscourse markers of each subcategory used by the CNWs with those of the EEWs. The numbers of interactive and interactional metadiscourse markers in RA's R&Ds by the CNWs and the EEWs were normalized to per 10,000 words (Table 4). To compare their salient pattern for a choice of metadiscourse markers, and to explore the metadiscourse repertoire of the CNWs, all subcategories and metadiscourse markers under each subcategory used by the two groups were ordered by frequency. Overall, our study found that the total frequencies (per 10,000 words) of metadiscourse markers used by the two groups were very close, and Chi-square analysis showed no significant difference between them ($p < 0.05$). However, a significant difference was found in both interactive and Interactional metadiscourse between groups. The CNWs used significant, more interactive metadiscourse markers than the EEWs but significantly fewer interactional ones ($p < 0.05$, Table 4). The results seem to indicate that, just as EFL writers, the CNWs attach greater importance to guiding readers through texts, but have less sensitivity to the need to involve readers into their texts and communicate with their readers. Both groups used interactive metadiscourse at a higher frequency than interactional ones. This could be attributed to the particularity of RAs in engineering disciplines: The major task of the author is to guide the reader through the texts before they can represent themselves to connect with their readers [7].

Table 4. Descriptive statistics for interactive and interactional metadiscourse of the two subcorpora.

Category	CNWs		EEWs		X^2	p
	Raw Frequency	Per 10,000 Words	Raw Frequency	Per 10,000 Words		
Interactive	2132	291.4	2198	253.3	21.8484	0.0000
Interactional	1096	149.6	1694	195.2	48.1108	0.0000
Total	3228	441.0	3892	448.6	0.5052	0.4772

4.1.1. Interactive Metadiscourse

In general, the CNWs used significantly more interactive metadiscourse markers in their RA's R&D than the EEWs (Table 5). Among the five categories, both groups used Transitions most frequently, followed by Endophorics, Frame Markers, Code Glosses, and Evidentials. Between the two groups, the CNWs used more Transitions and Endophorics, with a significant difference in Transitions, while the EEWs used the other three categories more frequently.

Table 5. Chi-square analyses of frequencies of interactive metadiscourse of the two subcorpora.

Interactive	CNWs			EEWs			X ²	p
	Raw Frequency	Per 10,000 Words	%	Raw Frequency	Per 10,000 Words	%		
Transitions	965	131.9	45.2	764	88.1	34.8	71.3485	0.0000
Endophorics	669	91.4	31.3	756	87.1	34.4	0.8343	0.3610
Frame Markers	236	32.3	11.1	290	33.4	13.2	0.1648	0.6848
Code Glosses	141	19.3	6.7	204	23.5	9.3	3.3067	0.0689
Evidentials	121	16.5	5.7	184	21.2	8.4	4.5442	0.0330
Total	2132	291.4	100	2198	253.3	100	21.8484	0.0000

Transitions

Transitions was the most frequent subcategory of interactive metadiscourse in both subcorpora, accounting for over 1/3 of the total interactive metadiscourse markers. Among the five categories of interactive metadiscourse, it was the only one that the CNWs used significantly more frequently than the EEWs ($p = 0.0000$). Compared with previous findings [7,8] which found that Chinese L1 writers of RAs normally used significantly fewer Transitions than English writers, our study found that the CNWs showed more preference for Transitions than the EEWs when writing English RA's R&Ds (Table 6). All three categories of Transitions saw higher occurrences in the CNW subcorpus than in the EEW one. Specifically, the CNWs used more than double the Additions that the EEWs used, especially for presenting results, while the EEWs preferred Contrasts, which were mainly used to compare their studies with other studies.

Table 6. Transitions used by the two groups (per 10,000 words).

Subcategories	CNWs			EEWs		
	F *	V **	Top 3 ***	F	V	Top 3
Addition	52.8	13	and (32.4) ****, moreover (4.0), in addition (3.4)	24.9	10	and (14.8), on the one/other hand (2.3), moreover (1.7)
Contrast	41.8	11	however (13.4), but (9.6), while (8.2)	32.5	9	however (11.6), while (6.3), although (5.5)
Consequence	36.4	9	therefore (10.1), thus (7.4), because (6.3)	2.5	10	thus (6.5), therefore (6.2), since (4.6)
Total	131.9	33		88.1	29	

Note: * per 10,000 words; ** variety; *** most frequently used three metadiscourse markers of each subcategory by each group; **** the token per 10,000 of the specific metadiscourse marker.

The CNWs seemed to be overdependent on explicit cohesive devices to realize discursive relations while neglecting the role of other discourse strategies, such as the logical arrangement of information, in building coherence. An extreme case of such overuse was that some CNWs redundantly used a paratactic or progressive transitions to lead every single sentence of a paragraph, as in Example 3.

Example 3. *The elastic-like property of PS/CeO₂ makes them deformed when stressed by the polishing pressure Meanwhile, as PS/CeO₂ can well adapt to pad asperities, a lower and more uniform contact stress is formed on the wafer surface, making the wafer removed gently and uniformly. Moreover, the lower contact stress will induce a lower mechanical damage on the wafer surface. Also, the buoyant effect of PS/CeO₂ is helpful to improve polishing quality, because it makes PS/CeO₂ difficult to precipitate and agglomerate. Therefore, PS/CeO₂ can contribute to better polishing quality compared with CeO₂. (CNW32).*

In Example 3, Additions such as *meanwhile*, *moreover*, and *also* were used as the starting point of every sentence in order to highlight the parallel or progressive relationship between adjacent sentences. However, such use seems awkward when the relationship between clauses is quite obvious or when other relationships are more appropriate, e.g., the cause-and-effect relationship between the first two sentences.

The variety of Transitions used by the CNWs was similar to that of the EEWs. However, some Transitions used by the CNWs seemed to be inappropriate. In Example 4, *whereas* is mistakenly used as an adverb rather than a conjunction to indicate a contrast. The CNWs also overused sentence initials *and*, *but*, and *so*, markers generally preferred in oral rather than written English, to show additive, contrastive, or causal relations between clauses. Such inappropriateness suggests the CNWs' inadequate knowledge of the usage of Transitions in academic genres.

Example 4. *For gap width, whereas, the most probable reason for the great influence on spark discharge energy was the volume of the plasma tunnel increases with the increase of gap width. (CNW27).*

Endophoric Markers

Compared with other sections of RAs, the R&D section normally sees a much higher density of Endophoric markers because they are frequently used in this section to direct readers towards visuals as well as facts, theories, methods, and research results presented before or after. In our study, Endophoric Markers had the second highest occurrences in both subcorpora, which showed similarity in both quantity and variety, despite some minor differences, in which the EEWs used slightly more non-linear Endophoric markers but fewer linear ones than the CNWs (Table 7). This result is similar to that of Lee and Casal's results [23] on the metadiscourse features in R&D chapters of English PhD dissertations. Both groups used a great number of non-linear Endophoric markers to refer to figures or tables. As argued by Hyland [3], writings in hard disciplines are "typically semiotic hybrids", and are thus characterised by a great density of Endophoric Markers to build connections between text and visuals. These markers were used especially when writers guided readers through the reporting or interpretation of results, providing concrete support to enhance argument validity (Example 5).

Table 7. Endophoric Markers used by the two groups.

Subcategories	CNWs		Top 3	EEWS		Top 3
	F	V		F	V	
Non-linear	77.5	6	in Table/Figure/Figure/Equation (47.4), ((see) Figure/Table/Figure) (20.2), from Table/Figure/Figure (4.8)	79.5	7	in Table/Figure/Figure/Equation (42.0), ((see) Figure/Table/Figure/Equation X*) (27.1), Table/Figure/Figure X (as the subject) (4.3)
Linear	13.9	6	above (5.7), following (5.1), in Section X (1.9)	7.6	6	above (2.3), in Section X (2.3), following (1.2)
Total	91.4	12		87.1	13	

Note: * X is used to stand for an Arabic number.

Example 5. Figure 7 shows the results of the experiments. . . . It was found that the SDE increase with the increase of discharge gap width, the SDE is 591 μJ with the gap width of 1.5 μm , 517 μJ with gap width of 3.5 μm and 948 μJ with the gap width of 6 μm as Figure 7a shows. (CNW27).

Both groups used Endophoric Markers with similar linguistic structures, with the prepositional phrase *in + visual* (e.g., in Figure 1) being the most frequent. However, the degree of preference was slightly different. While the EEWs used parenthesized structure (Figure X) much more frequently than the CNWs, the CNWs adopted an *as shown/seen* clause much more often when presenting the information of visuals. Among Linear Endophoric Markers, the EEWs used the retrospective *above* and prospective *following* far less than the CNWs, which is probably because the former preferred to precisely locate the mentioned materials (e.g., in Section X).

Frame Markers

The two subcorpora show similarity in both the frequency and variety of Frame Markers (Table 8). Congruent with Lee and Casal [23], Sequencers predominated in both subcorpora, followed by Announcers, while Discourse labels and Topicalisers appeared rarely.

Table 8. Frame Markers used by two groups.

Subcategories	CNWs			EEWS		
	F	V	Top 3	F	V	
Sequencers	22.4	11	then (7.5), finally (2.7), first (2.7)	21.2	12	then (18.2), numbering (4.0), finally (3.5)
Announcers	8.3	8	(in) this paper (3.4), three (1.4), here (1.2)	10.9	11	(in) this section/part (2.7), here (2.3), in this study (2.1)
Stage labels	1.4	4	in . . . step (0.5), in conclusion (0.5), at this stage (0.1)	0.9	3	in . . . step (0.6), at this stage (0.2), in summary (0.1)
Topicalisers	0.1	1	with regard to (0.1)	0.3	1	with regard to (0.3)
Total	32.3	24		88.1	27	

Although the CNWs used a similar number of Sequencers as the EEWs, they used the word *then* much more frequently when introducing the next item in a series of actions. In addition, the CNWs used much fewer Announcers than the EEWs. When reporting the aim or results of the study, the EEWs normally referred to *this study* (Example 6), while the CNWs often spoke of *this paper* (Example 7). *This paper* generally collocates with reporting verbs such as *discuss* to present the content of the RA. However, in the CNW data, it often inappropriately collocated with action verbs such as *use* to report the work done by the researchers.

Example 6. *This study is to investigate the structural changes occurring to hardwood Alcell TM lignin as a result of fiber devolatilization.* (EEW7).

Example 7. *In this paper, multiple linear regression method is used to study the correlation between operating temperature, wind speed, relative humidity and the coefficients k, n in the model.* (CNW25).

Code Glosses

Code Glosses were infrequent in both subcorpora, and no significant difference was found (Table 9). Between the two categories of Code Glosses, Reformulations were used only slightly more than Exemplifications, which is somewhat different from Lee and Casal's study [23] that found significantly more Reformulations than Exemplifications in their corpus. The two groups also showed similarity in a variety of expressions for the two

subcategories of Code Glosses. A difference between the two groups lies in the preferred Reformulation, with EEWs favoring the abbreviation *i.e.*, (Example 8) and the CNWs preferring *mean* (Example 9). While both expressions are an elaboration or explanations of ideas, the former seems more concise than the latter, indicating that the CNWs probably were less adequate in presenting information in a more concise way.

Table 9. Code Glosses used by two groups.

Subcategories	CNWs			EEWS		
	F	V	Top 3	F	V	Top 3
Reformulation	11.1	7	mean (3.8); <i>i.e.</i> , (3.6), namely (1.4)	12.4	9	<i>i.e.</i> , (6.5); or (1.2), that is (0.7)
Exemplification	8.2	4	such as (4.0); for example (1.5); for instance (1.1)	8.6	5	such as (4.6); for example (2.2); <i>e.g.</i> , (1.4)
Total	32.3	24		88.1	27	

Example 8. *This is also observed in this work, i.e., a decrease of the lattice spacing and an increase in orientation, crystallite size and pore diameter. (EEW6).*

Example 9. *PSNR of OMP is finite, which means there are some errors in reconstruction. (CNW10).*

Evidentials

Evidentials occurred infrequently in both subcorpora (Table 10). Evidentials are expected to occur in RA's R&Ds when writers compare their own findings with those of others or draw upon others' findings to interpret or support their own. However, unlike introductions where literature is densely referenced, R&Ds require relatively fewer references to literatures. The CNWs used fewer Evidentials than the EEWs, suggesting that they relied less on the studies of other researchers to support their arguments or made fewer comparisons with other studies.

Table 10. Types of Evidentials used by two groups.

Subcategories	CNWs	EEWs
	F	F
Integral	2.9	5.6
Non-integral	13.9	17.9
Total	16.5	23.5

Both groups used many more Non-integral Evidentials (a cited source within parentheses) than Integral Evidentials (a cited source incorporated as part of the statement), which is in line with Hyland's finding that non-integral citations are preferred in hard disciplines to give prominence to the research and less emphasis to the role of the researchers [2].

The CNWs used about half as many integral Evidentials as the EEWs. Integral Evidentials function to "foreground individual interpretations, alternative perspectives, and human agency in knowledge construction" and "allow writers to show their stance and make evaluations" [5] (p. 28). The low frequency of this type in the CNW subcorpus indicates that the CNWs intended to make their writing more objective and impersonal by concealing the role of evaluators. By using Non-integral Evidentials instead of Integral ones, the CNWs tended to "adopt a non-committal stance that acknowledges or distances themselves from cited sources" [46]. This could be attributed to the CNWs' "self-perceived peripheral status in the academic discipline, or their traditional Chinese values of collectivism" [36] (p. 15).

4.1.2. Interactional Metadiscourse

The CNWs used significantly fewer interactional metadiscourse markers than the EEWs (Table 11, $p = 0.0000$). This result echoes the findings of some previous studies [8], and suggests that the CNWs probably attend less to the writer–reader interaction than the EEWs. Among the five categories, the CNWs employed significantly fewer Hedges and Attitude Markers ($p = 0.0000$) but significantly more Self-mentions ($p = 0.0004$) than the EEWs.

Table 11. Chi-square analyses of frequencies of interactional metadiscourse by the CNWs and EEWs.

Interactional Metadiscourse	CNWs			EEWs			X^2	p
	F	Per 10,000 Words	%	F	Per 10,000 Words	%		
Hedges	441	60.2	40.0	745	85.9	44.0	35.3045	0.0000
Attitude Markers	244	33.3	22.3	507	58.4	30.0	53.4325	0.0000
Self-Mentions	158	21.6	14.5	123	14.2	7.3	12.4579	0.0004
Boosters	146	20.0	13.4	161	18.6	9.5	0.4059	0.5241
Engagement Markers	106	14.5	9.7	158	18.2	9.3	3.3362	0.0678
Total	1095	149.6	100	1694	195.2	100	48.1108	0.0000

Hedges

Hedges were the most frequent interactional metadiscourse and accounted for almost half of the total interactional metadiscourse in both subcorpora (Table 12). Indeed, Hedges tend to occur more frequently in R&D than in any other RA section because they facilitate the presenting of claims or arguments in a polite and cautious manner, opening up an agreeable space for the negotiation of alternative explanations of potential results, speculating on the limitations of the present study, and indicating the practical implications for future research [6]. However, the CNWs used far fewer hedges, less than two thirds of the amount used by the EEWs. In cases where it would be more appropriate to use Hedges to demonstrate tentativeness in presenting plausible reasoning, some CNWs made bare assertions, as in Example 10:

Table 12. Hedges used by the two groups.

Subcategories	CNWs			EEWS		
	F	V	Top 3	F	V	
Inference	20.2	11	indicate (12.6), suggest (3.3), assume (0.5)	25.1	14	indicate (11.1), suggest (8.0), seem (2.4)
Probability	15.0	9	may (4.5), could (3.7), would (2.9)	29.8	12	may (8.0), could (6.5), would (5.0)
Approximate limitation	9.7	8	mainly (4.0), almost (2.3), tend to (1.9)	10.5	12	typically (2.0), generally (1.5), mainly (1.4)
Approximate quantity	9.2	4	about (4.4), approximately (2.5), almost (1.2)	11.5	4	about (5.8), approximately (4.7), almost (0.7)
Approximate degree	4.7	4	relatively (3.3), nearly (1.2), somewhat (0.1)	6.6	3	relatively (4.7), nearly (1.5), almost (0.3)
Approximate frequency	1.4	2	usually (1.0), often (0.4)	2.3	4	often (1.4), usually (0.7)
Total	60.1	24		88.1	27	

Example 10. *The hydraulic conductivity predicted by the proposed model is more consistent with the experimental values, while the model by Watanabe overestimates the measured hydraulic conductivities. (CNW23).*

In Example 10, the use of *overestimates* without any down-toner or support for this claim makes the criticism of another researcher's work rather face-threatening. This shows that some CNWs were neither clearly aware of their relationship with readers in negotiating propositional information, nor familiar with the devices they could use to downplay their assertiveness and achieve politeness.

Both groups used the subcategories Probability (Example 11) and Inference (Example 12) predominantly, while Approximate degree and Approximate frequency were relatively infrequent. In R&Ds, writers tend to make speculations about the possible meaning or significance of their research findings and, at the same time, avoid direct personal responsibility for their statements.

Example 11. *The non-detection of a-SiC in the XRD results above may be attributed to two factors. (CNW2).*

Example 12. *This also can indicate that GASA can reconstruct noiseless signal accurately. (CNW16).*

The varieties of hedging expressions used by the two groups were similar, though the CNWs used slightly fewer varieties than the EEWs. Both groups mainly resorted to modal adverbs and lexical verbs to indicate cautiousness in making claims, followed by the use of modal verbs to express uncertainty, probability, or approximation. Adjectives and nouns were rarely used, especially by the CNWs. They also used a few compound-hedging structures (two or more hedges used together, e.g., *seem to indicate*) to increase the strength of tentativeness. In addition, various hedging expressions ranked similarly in terms of frequency in the two subcorpora, albeit with a few exceptions. Words such as *approximately, typically, likely, seem* were strongly preferred by the EEWs than by the CNWs, and a few advanced words, which were used by the EEWs, though only occasionally, such as *predominantly, postulate, indicative*, did not occur in the CNW subcorpus at all.

Additionally, the misuse of hedges occurred in the CNWs' data. A typical case was the modal verb *could*, which was often used in the phrase *it could be seen*, where *can* should be used to express ability. In Example 13, *could* contradictorily co-occurred with *clearly*, the former expressing uncertainty while the latter certainty. This suggests that the CNWs were probably not clear about the functions of some hedging expressions.

Example 13. *It could be clearly seen that the differences of coverage ratio among the different algorithms are decreasing. (CNW8).*

Attitude Markers

As Hyland has pointed out, RA writers frequently use Attitude Markers to express their judgments, evaluations, and views on textual information [47]. Both the CNWs and the EEWs used fairly large numbers of Attitude Markers in their R&Ds (Table 13). However, the EEWs used many more Attitude Markers than the CNWs, indicating that the latter were more conservative at explicitly marking personal attitudes in R&Ds. This is probably because the CNWs regarded Attitude Markers as signifying subjectivity, which may be discrepant from their perceived conventions of RAs [35]. In addition, compared with the EEWs, the CNWs used a much smaller range of Attitude Markers, indicating their limited repertoire of Attitude Markers.

Table 13. Attitude Markers used by the two groups.

Subcategories	CNWs			EEWs		
	F	V	Top 3	F	V	Top 3
Evaluative adjectives	18.2	24	good (2.7), important (2.5), superior (2.5)	26.0	44	important (5.5), good (3.2), best (1.6)
Attitude verbs	11.2	10	confirm (3.1), contribute (2.5), be consistent with (2.3)	18.3	17	expect (6.6), confirm (5.0), be in agreement with (2.5)
Obligation modals	2.9	2	should (2.3), must (0.5)	4.1	2	should (2.9), must (1.3)
Attitude nouns	1.6	4	advantage (0.8), superiority (0.5), disadvantage (0.1)	2.7	10	advantage (0.7), challenge (0.3), confidence (0.2)
Affective adverbs	1.1	2	interestingly (0.8), importantly (0.3)	2.9	9	interestingly (1.0) effectively (0.3), perfectly (0.3)
Total	33.3	42		58.4	82	

The EEWs used a much wider variety of Attitude Markers than the CNWs, especially Evaluative adjectives. This was the most common subcategory in both subcorpora, taking up about half of the total Attitude Markers, with *important* and *good* being most frequent (Example 14). Attitude verbs were the second preferred subcategory, while Obligation modals, Affective adverbs, and Attitude nouns appeared infrequently in both subcorpora. The verb *confirm* (Example 15) and the two phrases *be in consistent with* (Example 16) and *be in agreement with* were preferred by both groups to support their initial hypothesis or to compare their findings with previous studies. This shows the tendency for engineering writers to strengthen persuasiveness by showing how they themselves respond to the referenced material. However, the verb *expect* (Example 17), which topped the EEW's list when stating whether their results were in line with their own or readers' expectations, was far less frequently used by the CNWs, indicating that the CNWs were less ready to address their readers' perceptions or to get them involved in the reasoning processes. In addition, the verb *allow* (Example 18), which was preferred by the EEWs, often to denote a benefit of a finding, a material, etc., did not appear in the CNW's subcorpus at all, suggesting the CNWs' unfamiliarity with such usage of this word.

Example 14. *What is more important, convergence time of NPCGA increases very slowly and slightly with N , K or R increases. (CNW8).*

Example 15. *The experiments confirmed that the transition from the nano- to the coarse-grained SiC is extremely abrupt. (CNW2).*

Example 16. *On the whole, the above results are consistent with Theorem 1. (CNW21)*

Example 17. *This is expected due to the increase in number of filler aggregates within the scattering volume. (EEW5).*

Example 18. *This analysis allows the understanding of what actions must be undertaken to minimise the environmental impact of the process. (EEW31).*

Self-Mentions

Self-mention was the third most frequent interactional metadiscourse in both subcorpora (Table 14). Contrary to previous findings which showed that Chinese writers tended to avoid the use of Self-mentions [7], our corpus reveals a significantly more frequent use of Self-mentions by the CNWs than the EEWs. Self-mention is a promotional device which gives a community-approved persona and consolidates the writer's credibility among

other community members [29]. Both groups tended to use plural forms of the first-person pronoun to indicate the unity of the research group or the spirit of collectivism; no singular first-person pronoun was found in our corpus.

Table 14. Self-mentions used by the two groups.

Subcategories	CNWs	EEWs
First-person plural	21.4	14.2
we	(16.5)	(8.9)
our	(4.2)	(5.0)
us	(0.7)	(0.1)
ours	(0)	(0.2)
Third-person author	0.1	0
Total	21.6	14.2

An examination of the distribution of *we* by their main discursual functions reveals differences between the two groups (Table 15). In the EEWs subcorpus, *we* was mostly used to present an argument or speculation. In particular, it most frequently collocated with epistemic verbs of judgement, such as *assume* and *believe*, occupying about one-third of the instances *we* was used by the EEWs. *We*, for describing experimental procedures and for reporting findings, accounted for one-fourth respectively. In the CNW subcorpus, however, *we* mostly co-occurred with action verbs such as *compare*, *conduct*, *perform*, for describing experimental procedures (Example 19), which took up over half of the instances of *we* used by the CNWs. This indicates that the CNWs were more ready to assume the role of researchers than of evaluators. Additionally, *we* tended to appear frequently in such collocations as *we can see from*, *we can find*, mainly for reporting findings (Example 20), yet its co-occurrence with epistemic verbs was rather infrequent (Example 21).

Table 15. Distribution of *we* used by the two groups across functions.

Functions	CNWs		EEWs	
	F	%	F	%
Describing procedures	9.0	54.5	2.4	27.2
Reporting findings	3.8	23.1	2.4	27.2
Outlining discursual structure	0.8	4.9	0.7	7.8
Claiming common understanding	0.4	2.4	0.0	0
Defining or identifying	0.4	2.4	0.3	3.8
Presenting arguments or speculations	2.1	12.4	3.0	33.8
Total	16.5	100	8.9	100

Example 19. *In order to verify the effectiveness of the proposed RUIO we compare it with Jeang's PIO presented in [20].* (CNW12).

Example 20. *From the figure we can see that in the unfilled profiles there is a significant difference in heat retention between samples printed in the two orientations.* (CNW24).

Example 21. *We presume that more TiO₂ existed in native oxide films on \x{A6}\x{C1} phase due to its higher Ti content.* (CNW34).

Boosters

Both CNWs and EEWs used Boosters at a relatively low frequency (Table 16). As Hyland demonstrates, this is probably because the results in engineering disciplines are not that open for negotiation compared with those in the arts and humanities, and engineering writers tend not to use strong generalization so as to avoid refutation by readers [48]. The frequencies of Boosters used by the two groups were very close.

Table 16. Boosters used by the two groups.

Subcategories	CNWs			EEWS		
	F	V	Top 3	F	V	Top 3
Emphatic	17.2	19		15.9	16	
(Adverbs)	(13.1)	(12)	obviously (3.8), very (3.4), clearly (2.7)	(13.2)	(10)	very (6.5), clearly (2.8), always (1.6)
(Projected) clauses	(2.2)	(5)	it is clear that (0.8), it is evident that (0.5), there is no doubt that (0.4)	(3.5)	(4)	it is clear that (1.8), it is evident that (0.9), it is apparent that (0.6)
(Verbs)	(1.9)	(2)	demonstrate (1.6), believe (0.3)	(0.7)	(2)	demonstrate (0.6), believe (0.1)
Amplifying boosters	2.8	4	completely (1.9), always (0.5), totally (0.4)	2.7	3	always (1.6), completely (0.9), totally (0.1)
Total	20.0	25		18.6	19	

Between the two subcategories of Boosters, the frequency of Emphatic Boosters far outweighed Amplifying Boosters. Both groups resorted to adverbs, projected clauses, and verbs to intensify their claims, with adverbs taking up the great majority. However, the degree of their preference for various boosting expressions displayed some differences. For example, the adverb *obviously* was the most favored Booster by the CNWs (Example 22), yet it did not occur in the EEWS' data. The CNWs used less projected structure such as *it is clear that* to show confidence in their arguments than the EEWS.

Example 22. *Obviously, the composite coatings prepared by conventional methods mainly focus on corrosion resistance, whereas electrical insulation is usually ignored with few published literatures. (CNW44).*

Engagement Markers

Engagement Markers are mainly employed to connect readers and attract their attention to the writers' findings or statements [34]. In this study, both groups used Engagement Markers in their RAs sparingly (Table 17). The CNWs used fewer Engagement Markers than the EEWS, indicating that the former were probably less ready to create dialogic space with readers. Among the five subcategories of Engagement Markers, Directives were the main strategy adopted by both groups to engage readers, followed by some shared knowledge references and a few reader mentions; neither used Questions or Personal asides. Within the subcategory of Directives, the CNWs used more Predicative structures than Imperatives, while the EEWS used more Imperatives. The imperatives *see* and *note* predominated the EEWS' list (Example 23), especially when navigating readers towards tables or figures, while it was used only occasionally by the CNWs, who favored the predictive structure *it can been seen* (Example 24). Since predictive structures are somewhat less explicit than imperatives in initiating reader participation, our findings seem to suggest that the CNWs tended to keep a certain distance with readers when presenting results or arguments.

Table 17. Engagement Markers used by the two groups.

Subcategories	CNWs			EEWS		
	F	V	Top 3	F	V	Top 3
Directives	10.1	15		15.7	12	
(Predicative structure)	(6.2)	(9)	it is/can be seen (4.0), it should be noted that (0.8), it is worth noting that (0.3)	(6.8)	(9)	it should/can be noted that . . . (3.0), it is/can be seen that (2.1), it is . . . to note that (0.3)
(Imperatives)	(4.0)	(6)	note (1.0), see (1.0), suppose (0.7)	(8.9)	(3)	see (5.2), note (3.3), consider (0.3)
Share knowledge references	2.6	4	as can be seen (0.8), it is known that (0.7), as is known (0.5)	3.0	4	as seen (1.4), it is (well) known (0.7), be known to (0.7)
Reader mentions	1.8	2	(as) we can see (1.1), one can see (0.7)	0.7	3	one (can) see/notice/conclude (0.7)
Total	14.5	21		19.4	19	

Example 23. *Note that after decoding one frame, the global UTC time is available and therefore the timing module can be safely powered off. (CNW5).*

Example 24. *It can be seen from Figure 9 that the volume of pores decreases with increasing amount of OSRA, which is consistent with the result of total porosity in Table 9. (CNW20).*

4.2. Findings of the Questionnaire Survey

To answer the second research question, i.e., whether the Chinese novice writers' choice of metadiscourse is a result of their perception of RA writing, we conducted a questionnaire survey on the CNWs' perception of metadiscourse use in RA's R&Ds.

4.2.1. Functions of RA's R&D

We asked the respondents to give their ideas on the major functions of RA's R&Ds. Almost all the respondents reported that R&Ds should include a brief summary of results, an interpretation of results, and directions of future study. It reveals that the CNWs were aware of the major purpose of this part-genre in interpreting their results so as to reveal the answers to research questions. However, most respondents failed to recognize other important functions of RA's R&D, such as the negotiating of knowledge claims in the context of the published knowledge and demonstrating the value of the study. Only about 20% of the respondents mentioned these functions. Thus, the respondents seemed to have an incomplete knowledge of the functions of RA's R&Ds, especially regarding its role in negotiating different views and claiming novelty. This may limit their use of some categories of metadiscourse, such as Evidentials and Attitude Markers.

4.2.2. Awareness of Writer–Reader Interactions

Writer–reader interaction emphasizes the dialogic nature of a text, meaning that writers should anticipate readers' questions or reactions to what is written and address such reactions accordingly [49]. To explore the CNWs' awareness of writer–reader interactions, we asked whether they would anticipate the potential readers of their RAs. About two-thirds of the respondents reported that they would never or rarely, saying that they simply tried to imitate RAs written by experts but did not think how they should deploy the language to meet readers' expectations. This reveals their lack of awareness of RA writing as a kind of social interaction, wherein knowledge is mutually constructed, negotiated, and created. This also indicates their traditional view of RAs as a passive expression of knowledge which requires little use of rhetorical strategies.

The respondents were also asked how they would revise their RAs in anticipation of readers' needs. About three-fourths of the respondents said that they would make

their writing more easily understood, mainly by elaborating results, increasing coherence, defining jargons, and attending the accuracy of information. This reveals their efforts to facilitate readers' comprehension by increasing the clarity of information presentation, which further shows their emphasis on propositional ideas and their readiness to use interactive metadiscourse in their writing. However, their responses revealed little awareness of adjusting the interpersonal relationship with readers, e.g., the choice of personal pronouns to adjust the writer–reader distance, the use of persuasive strategies to convince readers of the research's validity, or the use of an appropriate tone to make readers emotionally comfortable. Thus, it seems that they were not fully aware of the dialogic strategies for inviting or restricting reader participation.

4.2.3. Perception of Metadiscourse Use

The respondents were asked how often they thought each category of metadiscourse should be used in RA's R&Ds. Table 18 summarizes the mean score of their perceived frequency for each category, and Figure 1 shows the distribution of the CNWs' perceived frequency of using each category.

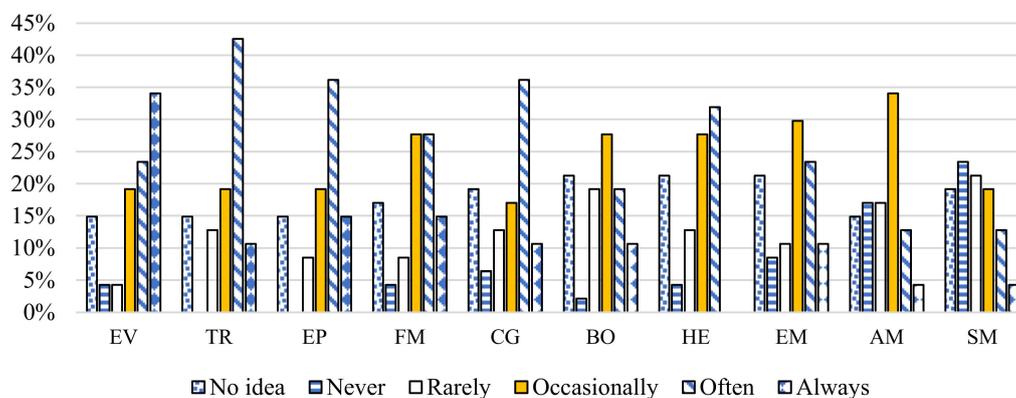


Figure 1. CNWs' perceived frequency of using each category of metadiscourse.

Table 18. Mean score of their perceived frequency for each category.

	EV *	EP	TR	FM	CG	BO	HE	EM	AM	SM
Mean	3.34	3.13	3.06	2.89	2.70	2.53	2.45	2.45	2.26	1.96
SD	1.76	1.56	1.54	1.65	1.67	1.63	1.52	1.61	1.41	1.46

* EV: Evidentials; EP: Endophorics; TR: Transitions; FM: Frame Markers; CG: Code Glosses; BO: Boosters; HE: Hedges; EM: Engagement Markers; AM: Attitude Markers; SM: Self-mentions.

For each category, about 15–21% of the respondents had no idea whether they should use it in R&Ds. This shows that a considerable proportion of the CNWs were not clear about the functions of metadiscourse, or that their choice of metadiscourse was an unconscious behavior. The categories of interactive metadiscourse generally received higher mean scores than interactional metadiscourse, indicating that the CNWs recognized the need to use more interactive markers than interactional ones in R&Ds. Among the 10 subcategories, Evidentials obtained the highest mean score (3.34), while Self-mentions the lowest (1.96).

The respondents were also asked to give reasons for their perception. For interactive categories, especially Evidentials, Transition, and Endophorics, most respondents chose "often" or "always" regarding the frequency of their use. The main reasons the CNWs gave for their frequent use of Evidentials were to avoid plagiarism, maintain academic integrity, and to gain credibility, as seen, for example, in Response 1:

Response 1. *It's very important to provide the source of a citation. This is the rule of scientific writing. Otherwise, it will be considered as plagiarism.*

This shows the CNWs generally have an awareness of academic ethics. The reasons for using Transitions and Frame Markers are quite similar, mainly to increase logicity, coherence, and organization. For example, Response 2 noted:

Response 2. *Research articles should be in logical order. These expressions (Transitions) can help writer express his views logically and make the writing more organized.*

Interestingly, A few respondents also referred to the difference between English and Chinese when giving reasons for frequent use of Transitions (Response 3):

Response 3. *I want to make the writing more logical and organized. I learned in my English classes that English writing usually uses more connecting words to show logical relationship between sentences than Chinese writing.*

Note: The responses were originally in Chinese and were translated by the authors.

The main reason for using Endophorics was also related to the coherence of the text, as most respondents put it, and as shown in Response 4:

Response 4. *These expressions (Endophorics) can increase the connectiveness between various parts of the text.*

A few respondents took readers into consideration, saying that using these markers could make it easier for readers to follow, such as in Response 5:

Response 5. *In writing, these expressions (Frame Markers) can give readers a logical order, so that they can well grasp the author's thoughts and intentions.*

In terms of interactional metadiscourse, most respondents expressed a reservation in using them. The main reason for their avoidance of using Self-mentions was that these markers might increase the subjectivity of claims and reduce professionalism of the study, as shown in Response 6:

Response 6. *Of course I will never use Self-mentions when writing Results and Discussion. These expressions may highlight the role of authors and make the writing subjective. It is quite inappropriate to use them in RAs since we should keep neutral towards our results. I always use passive voice to avoid the use of first-person pronouns.*

When commenting on the use of Attitude Markers (Response 7) and Boosters (Response 8), respondents gave similar perceptions, saying that these expressions explicitly emphasized the writer's personal attitude and were inconsistent with the RA's feature of objectivity. For Hedges, most respondents believed that such expressions were seldom used in scientific writing because they would reduce the credibility of their conclusions, as in Response 9. Some believed that it was inappropriate or even wrong to employ expressions with emotional tendencies in their RAs because their supervisors did not encourage them to do that, as seen in Response 10.

Response 7. *The emphasis of Results and Discussion should be on the analysis of results, so I would be cautious in using these expressions so as not to reveal my personal inclination.*

Response 8. *Using these words (Boosters) does not conform to RA conventions. They either make my claims sound unreliable or make me appear overconfident.*

Response 9. *I would not use these words (Hedges) because they reduce the certainty of my findings and thus reduce rigorousness of reasoning.*

Response 10. *I learned from my supervisors that we should report our findings as objectively as possible when writing research articles. Thus, we should not express my attitude or emotion too often.*

Note: The words in parentheses are added by the authors.

A few respondents attributed avoidance of interactional metadiscourse to their novice writer-researcher identity, as shown in Response 11:

Response 11. *As a novice writer, I do not have the confidence in strengthening my argument. Only experts have the authority to do that.*

To further reveal the CNWs' general perception of the writing of the RA's R&D, we conducted thematic analyses to identify common ideas that emerged repeatedly in their responses. For the interactive metadiscourse, the most frequent words occurring in their responses included "objective" and "logic", showing that the CNWs considered the RA's R&D section as a genre of objectiveness and logicity that does not need much embellishment. This probably explains the exceptionally high occurrence of Transitions in the CNWs' corpus, which functions to facilitate the logical connection between sentences. For interactional metadiscourse, the most frequent words in their responses included "rigorous", "accurate", and "objective". This again shows that the CNWs probably attached a greater importance to objectivity compared to other characteristics of RAs.

5. Discussion: Metadiscourse Features and Identity Construction

Based on the results of the corpus study and the questionnaire survey, we can answer the third research question, i.e., how the metadiscourse choice of these novice writer-researchers reflects the construction of their writer identity, especially their discorsal self and authorial self.

5.1. Discorsal Self

Discorsal self refers to the intentional or unintentional self-representation of writers to claim membership to a community they feel affiliated with [19]. The questionnaire survey reveals the CNWs' awareness of RA as a specific discorsal genre, which requires the observation of specific conventions in order to be accepted by the research discourse communities. The CNWs used significantly fewer interactional metadiscourse markers than the EEWs. One reason, as revealed by the questionnaire responses, may be that the CNWs attach less importance to such interpersonal functions of RAs, such as recognizing, constructing, and negotiating social relations, realized by metadiscourse markers, than to propositional content. Among all kinds of metadiscourse, the CNWs used significantly more Transitions, but significantly fewer Hedges and less of other types of markers on attitude and stance than the EEWs, indicating their tendency to focus more on logicity and organization than interaction. This is mainly due to their perception of RAs as a genre of representations of objective reality and their perception of themselves as a reporter of their research and guide and navigator of the text rather than an evaluator of various views or as a creative originator. They also fail to see themselves as a discourse constructor and promoter who can manipulate rhetorical strategies to enhance the credibility of their research. This perception may arise from the reinforced instruction that natural sciences are characteristic of hierarchical knowledge structures, which gives rise to "an explicit, coherent, systematically principled and hierarchical organization of knowledge" [50] (p. 172, cited in [5]).

Their reserved use of interactional metadiscourse also reflects their view of RAs as a closed, fixed monologue by the writer himself/herself about the objective world, which requires little to deploy various genre options to engage readers. Failing to see the flexibility of disciplinary conventions, they feel more comfortable strictly conforming to the fixed sets of rules they mostly learn from EFL classes or from their supervisors. To make their writing sound more impersonal and objective, they purposefully reserve their own attitude in their RAs, and follow other conventions to achieve objectivity, which have been repeatedly reinforced by their EFL teachers or supervisors and which they consider as the most secure way to be recognized by the disciplinary community. In fact, many CNWs fail to recognize RA writing as a process in which knowledge is socially constructed through the deployment of rhetorical strategies to make their RAs both understandable and persuasive. Some CNWs even go so far as to see no point of arguing for the novelty and significance of their research, for the reason that this can be self-evident if new findings are presented objectively.

5.2. Authorial Self

The authorial self is mainly associated with writers' willingness to make claims and/or their reliance on external authorities to support those claims [19]. The questionnaire survey reveals that many CNWs reported to avoid using interactional metadiscourse markers, such as attitude markers and boosters to explicitly present their position, indicating their reluctance to deploy these resources to strengthen epistemic conviction. However, it should be noted that, although hard disciplines, compared with soft disciplines, may depend more on "procedural adequacy and methodological rigor" to achieve empirical authority than on personal voice or authority [6] (p.20), appropriate use of these metadiscourse resources will accentuate the persuasiveness and authoritativeness of the claims.

The authorial self is most typically represented by the use of first-person pronouns [45]. The high-frequency of Self-mentions in the CNWs' subcorpus apparently indicates that they did not spare to present a personal ethos in their writings. However, compared with the EEWs, who used the pronoun *we* mainly to present arguments and claims, the CNWs used *we* mainly for reporting factual information, such as procedures or findings, with which they had more confidence to present, though infrequently used them for interpreting results or presenting opinions, which carries more risk of being refuted. This agrees with some previous findings that Chinese EFL writers are reluctant to stake out a firm authorial identity in their RA's R&Ds [35,44]. This is further confirmed by the questionnaire results, which reveal that the CNWs tend to avoid first-person pronouns, mainly to mitigate the subjective voice of researchers and maximize the objectivity of the text.

Our results on the distribution of Self-mentions in the two subcorpora also reveals the different degrees of authority the two groups assumed for their research. One factor contributing to such a difference may be their different preferred rhetorical moves. Most the CNWs began their R&Ds with a summary of procedures, while the EEWs often began by directly with reporting results. In addition, the CNWs' R&Ds focused more on reporting results and interpreting results, giving little room for making an evaluation of related studies or for claiming the novelty of the present studies. This corroborates with Wu's research, which revealed that Chinese RA writers adopted a plain reporting of the research procedure and results as the main writing style [31]. This further shows that the CNWs tend to give much more prominence to the role of recounter of research over that of an evaluator or originator.

As inexperienced EFL writer-researchers, the CNWs' deficiency in both linguistic competence and disciplinary expertise may prevent them from taking more authoritative roles. Due to limited linguistic repertoire, they mainly resort to an imitation of experts' writings, especially when recounting methodology and reporting results, which requires a relatively low degree of language creativity. However, for evaluations and opinions, which require more originality of language, the CNWs often have to sacrifice the length and depth of those moves to reduce language mistakes. As Rafoth has suggested, inadequate language proficiency of EFL writers may hinder them from expressing their novel ideas and from constructing an identity as a confident researcher [51]. In addition, due to a lack of confidence in their disciplinary knowledge, they often hesitate to take an authoritative stance with which to comment on others' and their own studies, thus intending to hide their authorial identity and adopting a remote stance.

As Ivanič has pointed out, the construction of writer identity is both a conscious and subconscious choice of many possibilities of selfhood available to writers in the social context of writing [19]. In this process, a writer needs to deal with relationship between proximity, "relationship between self and community", and positioning, "the relationship between the speaker and what is being said" [48] (p. 36). A successful writer is able to locate an appropriate place where he/she "proclaims both individuality and membership of a group and a culture" (Ibid). For the CNWs, however, their perceived disciplinary conventions seem to prevent them from positioning themselves as competent researchers. Thus, when trying to achieve disciplinary proximity, they sacrifice their authority for objectivity. Their identity construction seems to lack the flexibility of choice, constrained by

their limited genre knowledge, and their attempts to align with their perceived prescribed disciplinary norms, which are reinforced through EFL education. It is also constrained by their unawareness of reader–writer interactions, in which their EFL education often neglects. It is important for EFL students to change the one-dimensional perception of RAs as an objective and detached writing and understand how proximity can serve as a resource that enables individual positioning, and how positioning themselves as a more authoritative writer-researcher can help them gain credits in the disciplinary community [48].

6. Conclusions

The present study investigates metadiscourse features of RA results and discussion written by Chinese PhD students through a corpus-based approach. Although no significant difference was found in the total frequency of metadiscourse markers between the CNWs and the EEWs, the CNWs were found to use significantly more interactive metadiscourse, but significantly fewer interactional metadiscourse. In addition, various degrees of difference were found in the frequency of specific categories and subcategories of metadiscourse. Our findings indicate that Chinese novice writer-researchers tend to attach more importance to organization and clarity of ideas than to the persuasiveness of their arguments. In addition, the varieties of the metadiscourse markers of each subcategory used by the CNWs were generally smaller than those of the EEWs, showing that the CNWs may have a limited repertoire of metadiscourse resources. The CNWs' perception of metadiscourse function in this part-genre was also studied through a questionnaire survey to see the influence of such perceptions on their practice of metadiscourse choice and their writer identity as reflected through such choices. The questionnaire survey reveals that these writers generally view RA as an organized, objective genre, thus trying to achieve logicity in the text as a recounter and reporter of research while remaining conservative in giving their authorial voice as an evaluator or knowledge creator.

This study shows that metadiscourse choice is somewhat influenced by the writer's perception of generic conventions, reader–writer interaction, and metadiscourse functions. EAP/ESP writing instruction plays an important role in developing EFL writers' appropriate perception. At present, EAP/ESP writing instruction in China has begun to adopt a genre-based pedagogy, focusing on the macro-structure of the research genre and rhetorical moves in various sections. Micro-level instruction mainly concerns typical sentence patterns and the general grammatical rules of the research genre. Such rules often leave learners with the impression that research writing is a rather fixed act, rather than a flexible process that allows individual choice for discourse construction. In addition, while much attention is placed on how to improve the accuracy of expressions, little is taught about how to achieve effective communication with readers, e.g., relationships between lexico-grammatical features and their rhetorical purposes. The study points to the need to place greater importance on salient linguistic features, e.g., metadiscourse, in the EAP/ESP curriculum. For instance, the functions of different categories of metadiscourse and linguistic resources to achieve these functions can be explored by studying authentic RA examples.

By comparing CNWs' metadiscourse features with those of the EEWs, we do not intend to suggest that EFL novice writer-researchers should strictly follow the practice of L1 English expert writers. Our aim is to show that novice writers should be assisted in developing their awareness of rhetorical purposes and strategies to facilitate the sustainable development of their writing competence. Corpus and discourse analysis approaches can be adopted in teaching to help learners understand possible metadiscourse patterns in RAs of different disciplines. Such approaches also increase their awareness of certain lexico-grammatical features that are not a result of fixed rules but the preference of a disciplinary community or the choice by the writer to achieve certain rhetorical purposes. EAP/ESP teachers should help EFL learners change the one-dimensional perception of RAs as an objective and detached writing and help them to understand written texts as writer–reader interaction. In addition, teachers should also equip learners with metadiscursive strategies

to express their professional opinions in RAs and establish their ethos as experts in their research areas.

The present study has some limitations, though. Firstly, the size of the corpus was relatively small because the RA's R&Ds in our language center database that fit our selection criteria were limited. Future studies will be based on a larger corpus of different disciplines so as to gain more generalizable findings. Secondly, the questionnaire survey was conducted only among Chinese novice writers. Future research may consider inquiring expert writers to investigate how they perceive the use of these rhetorical resources. Oral interviews could be adopted to elicit more salient responses from participants.

Despite its limitation, the present study contributes to the understanding of metadiscourse features of Chinese EFL writer-researchers and their relationships with writer identity construction. It also obtains lists of the metadiscourse markers of various subcategories and the writers' degrees of preference for them, providing EFP instructors and RA writers with useful references for metadiscourse choice. By adopting a questionnaire survey as well as quantitative and qualitative analyses of discourse, the study offers an explanation of metadiscourse choice by Chinese novice writer-researchers, indicating the benefits of using combined methods to conduct studies on discourse.

Author Contributions: Conceptualization, X.G.; Data curation, X.G. and Z.X.; Formal analysis, X.G. and Z.X.; Funding acquisition, X.G.; Investigation, X.G.; Methodology, X.G. and Z.X.; Project administration, X.G.; Resources, X.G. and Z.X.; Software, Z.X.; Supervision, X.G.; Validation, X.G.; Visualization, X.G.; Writing—original draft, X.G., Z.X.; Writing—review and editing, X.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Education Science “Fourteenth Five” Planning Project of Heilongjiang Province, China (Project No. GJB1421043); the Education Reform Project of Harbin Institute of Technology, China (Project No. 21MS044); the Philosophy and Social Science Prosperity Project of Harbin Institute of Technology, China (Project No. HIT.HSS.202128), and Mixed Teaching Mode Reform Project of Harbin Institute of Technology, China (2020.7).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the School of International Studies, Harbin Institute of Technology.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The questionnaire data presented in this study are available on request from the corresponding author. The corpus data were not publicly available since some of the investigated research articles in the data have not been published yet.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Morton, J.; Storch, N. Developing an authorial voice in Ph.D. multilingual student writing: The reader's perspective. *J. Second Lang. Writ.* **2019**, *43*, 15–23. [\[CrossRef\]](#)
2. Hyland, K. *Metadiscourse: Exploring Interaction in Writing*; A & C Black: London, UK, 2005.
3. Hyland, K. Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *J. Second Lang. Writ.* **2004**, *13*, 133–151. [\[CrossRef\]](#)
4. Hyland, K. Persuasion and context: The pragmatics of academic metadiscourse. *J. Pragmat.* **1998**, *30*, 437–455. [\[CrossRef\]](#)
5. Cao, F.; Hu, G. Interactive metadiscourse in research articles: A comparative study of paradigmatic and disciplinary influences. *J. Pragmat.* **2014**, *66*, 15–31. [\[CrossRef\]](#)
6. Hu, G.; Cao, F. Disciplinary and paradigmatic influences on interactional metadiscourse in research articles. *Engl. Specif. Purp.* **2015**, *39*, 12–25. [\[CrossRef\]](#)
7. Mu, C.; Zhang, L.J.; Ehrich, J.; Hong, H. The use of metadiscourse for knowledge construction in Chinese and English research articles. *J. Engl. Acad. Purp.* **2015**, *20*, 135–148. [\[CrossRef\]](#)
8. Li, Z.; Xu, J. Reflexive metadiscourse in Chinese and English sociology research article introductions and discussions. *J. Pragmat.* **2020**, *4*, 47–59. [\[CrossRef\]](#)
9. Jiang, F.; Hyland, F. Metadiscursive nouns: Interaction and cohesion in abstract moves. *Engl. Specif. Purp.* **2017**, *46*, 1–14. [\[CrossRef\]](#)
10. Poole, R.; Gnann, A.; Hahn-Powell, G. Epistemic stance and the construction of knowledge in science writing: A diachronic corpus study. *J. Engl. Acad. Purp.* **2019**, *42*, 1–11. [\[CrossRef\]](#)

11. Kawase, T. Metadiscourse in the introductions of PhD theses and research articles. *J. Engl. Acad. Purp.* **2015**, *20*, 114–124. [[CrossRef](#)]
12. Geng, Y.; Warton, S. Evaluative language in discussion sections of doctoral theses: Similarities and differences between L1 Chinese and L1 English writers. *J. Engl. Acad. Purp.* **2016**, *22*, 80–91.
13. Ament, J.; P'ares, J.B.; P'erez-Vidal, C. A study on the functional uses of textual pragmatic markers by native speakers and English-medium instruction learners. *J. Pragmat.* **2020**, *156*, 41–53. [[CrossRef](#)]
14. El-Dakhs, D. Variation of metadiscourse in L2 writing: Focus on language proficiency and learning context. *Ampersand* **2020**, *7*, 100069. [[CrossRef](#)]
15. Can, T.; Cangir, H. A corpus-assisted comparative analysis of self-mention markers in doctoral dissertations of literary studies written in Turkey and the UK. *J. Engl. Acad. Purp.* **2019**, *42*, 1–14. [[CrossRef](#)]
16. Liu, Y.; Buckingham, L. The schematic structure of discussion sections in applied linguistics and the distribution of metadiscourse markers. *J. Engl. Acad. Purp.* **2018**, *34*, 97–109. [[CrossRef](#)]
17. Sheldon, E. Knowledge construction of discussion/conclusion sections of research articles written by English L1 and L2 and Castilian Spanish L1 writers. *J. Engl. Acad. Purp.* **2019**, *37*, 1–10. [[CrossRef](#)]
18. Swales, J. *Research Genres: Explorations and Applications*; Cambridge University Press: Cambridge, UK, 2004.
19. Ivanič, R. *Writing and Identity: The Discoursal Construction of Identity in Academic Writing*; John Benjamins Publishing: Amsterdam, The Netherlands, 1998.
20. Dressen-Hammouda, D. From novice to disciplinary expert: Disciplinary identity and genre mastery. *Engl. Specif. Purp.* **2008**, *27*, 252. [[CrossRef](#)]
21. Zhao, C.G. Writer background and voice construction in L2 writing. *J. Engl. Acad. Purposes* **2019**, *37*, 117–126. [[CrossRef](#)]
22. Zhang, S.; Yu, H.; Zhang, L.J. Understanding the sustainable growth of EFL students' writing skills: Differences between novice and expert writers in their use of lexical bundles in academic writing. *Sustainability* **2021**, *13*, 5553. [[CrossRef](#)]
23. Lee, J.J.; Casal, J.E. Metadiscourse in results and discussion chapters: A cross-linguistic analysis of English and Spanish thesis writers in engineering. *System* **2014**, *46*, 39–54. [[CrossRef](#)]
24. Abdollahzadeh, E. Poring over the findings: Interpersonal authorial engagement in applied linguistics papers. *J. Pragmat.* **2011**, *43*, 288–297. [[CrossRef](#)]
25. Swales, J.M. *Genre Analysis*; Cambridge University Press: Cambridge, UK, 1990.
26. Bunton, D. Genre moves in Ph.D. thesis introductions. In *Academic Discourse*; Flowerdew, J., Ed.; Longman: Harlow, UK, 2002; pp. 57–75.
27. Dahl, T. Textual metadiscourse in research articles: A marker of national culture or of academic discipline. *J. Pragmat.* **2004**, *36*, 1807–1825. [[CrossRef](#)]
28. Hyland, K. Boosting, hedging and the negotiation of academic knowledge. *Text* **1998**, *18*, 349–382. [[CrossRef](#)]
29. Harwood, N. "Nowhere has anyone attempted . . . In this article I aim to do just that": A corpus-based study of self-promotional I and we in academic writing across four disciplines. *J. Pragmat.* **2005**, *37*, 1207–1231. [[CrossRef](#)]
30. Lafuente-Millán, E. Reader engagement across cultures, languages and contexts of publication in business research articles. *Int. J. Appl. Linguist.* **2013**, *24*, 201–223. [[CrossRef](#)]
31. Wu, G.Q. *Identity Construction in English and Chinese Research Articles: A Contrastive Study*; Zhejiang University Press: Hangzhou, China, 2013.
32. Wingate, U. 'Argument!': Helping students understand what essay writing is about. *J. Engl. Acad. Purp.* **2012**, *11*, 145–154. [[CrossRef](#)]
33. Aull, L.L.; Lancaster, Z. Linguistic markers of stance in early and advanced academic writing: A corpus-based comparison. *Writ. Commun.* **2014**, *31*, 151–183.
34. Hyland, K. Representing readers in writing: Student and expert practices. *Linguist. Educ.* **2005**, *16*, 363–377. [[CrossRef](#)]
35. Lee, J.J.; Deakin, L. Interactions in L1 and L2 undergraduate student writing: Interactional metadiscourse in successful and less-successful argumentative essays. *J. Second Lang. Writ.* **2016**, *33*, 21–34. [[CrossRef](#)]
36. Zhang, L.; Zhang, L.J. Fostering stance-taking as a sustainable goal in developing EFL students' academic writing skills: Exploring the effects of explicit instruction on academic writing skills and stance deployment. *Sustainability* **2021**, *13*, 4270. [[CrossRef](#)]
37. Zhang, F.; Zhan, J. Understanding voice in Chinese students' English writing. *J. Engl. Acad. Purp.* **2020**, *45*, 1–9. [[CrossRef](#)]
38. Intaraprawat, P.; Steffensen, M.S. The use of metadiscourse in good and poor ESL essays. *J. Second Lang. Writ.* **1995**, *4*, 253–272. [[CrossRef](#)]
39. Wu, S.M. Creating a contrastive rhetorical stance: Investigating the strategy of problematization in students' argumentation. *RELC J.* **2007**, *37*, 329–353.
40. Park, S.; Oh, S. Korean EFL learners' metadiscourse use as an index of L2 writing proficiency. *SNU J. Educ. Res.* **2018**, *27*, 65–89.
41. Bax, S.; Nakatsuhara, F.; Waller, D. Researching L2 writers' use of metadiscourse markers at intermediate and advanced levels. *System* **2019**, *83*, 79–95. [[CrossRef](#)]
42. Noble, W. Understanding metadiscoursal use: Lessons from a 'local' corpus of learner academic writing. *Nord. J. Engl. Stud.* **2010**, *9*, 145–169. [[CrossRef](#)]
43. Hyland, K. Options of identity in academic writing. *ELT J.* **2002**, *56*, 351–358. [[CrossRef](#)]
44. Hyland, K. Authority and invisibility: Authorial identity in academic writing. *J. Pragmat.* **2002**, *34*, 1091–1112. [[CrossRef](#)]

45. Isik-Tas, E. Authorial identity in Turkish language and English language research articles in sociology: The role of publication context in academic writers' discourse choices. *J. Engl. Specif. Purp.* **2018**, *49*, 26–38. [[CrossRef](#)]
46. Lee, J.J.; Hitchcock, C.; Elliott Casal, J. Citation practices of L2 university students in first-year writing: Form, function, and stance. *J. Engl. Specif. Purp.* **2018**, *33*, 1–11. [[CrossRef](#)]
47. Hyland, K. Stance and engagement: A model of interaction in academic discourse. *Discourse Stud.* **2005**, *7*, 173–192. [[CrossRef](#)]
48. Hyland, K. Genre, discipline and identity. *J. Engl. Acad. Purp.* **2015**, *19*, 32–43. [[CrossRef](#)]
49. Bakhtin, M. *The Dialogical Imagination*; University of Texas Press: Austin, TX, USA, 1981.
50. Bernstein, B. *Pedagogy, Symbolic Control and Identity: Theory, Research, Critique*; Taylor & Francis: London, UK, 1996.
51. Rafoth, B. *Multilingual Writers and Writing Centers*; Utah State University Press: Logan, UT, USA, 2015.