# The Use of Interactive Metadiscourse in Research Articles: A Corpus-Based Case Study of Astronautics

Yue Sun

Northwestern Polytechnical University, Xi'an, China
Email: syestherclemence@outlook.com

*Abstract*—Based on Hyland's interpersonal model of interactive metadiscourse, this study adopts astronautics as a research case to investigate the textual features of research articles, aiming to investigate the use of metadiscourse in research articles and assist academic writing for students who major in astronautics. A corpus consisting of 43 research articles published in *Journal of Guidance, Control and Dynamics (JGCD)* in 2020 and 2021 was built. AntConc 3.5.9 was utilized for searching and counting the standardized frequency results, so as to identify the most frequently used expressions. Results show that transition markers are most frequently used (35.72 per 1,000-word tokens), accounting for 65.62% of the total amount of interactive metadiscourse. The frequency of evidentials is the lowest, i.e., 0.26 per 1,000-word tokens, accounting for 0.48%. These indicate that transition markers play an important role in research articles in astronautics owing to the high requirement of logical expression. The low frequency of evidentials reflects low-level of dependency from others' opinions, focusing on own experiments and argumentations. The relatively high frequencies of frame markers and endophoric markers confirm that straightforward logic and clear structure together with transparent figures are preferred in research articles in astronautics.

*Keywords*—astronautics, interactive metadiscourse, self-built corpus

## I. INTRODUCTION

According to the press conference for *Blue Book of China Aerospace Science and Technology Activities (2021)*, 2021 is an extraordinary year in the history of human spaceflight, with a total of 146 rocket launches around the world, the largest number of launches in history. This achievement is even more remarkable considering the fact that these were achieved in the post-pandemic era. Therefore, it is relatively significant to investigate research articles in astronautics and summarize the textual features, exploring possible tactics in order to assist academic writing of this field.

This study adopts astronautics as a research case to investigate the textual features of research articles. Adopting Hyland's interpersonal model of interactive metadiscourse [1] as the theoretical framework, this study aims at investigating the use of metadiscourse in research articles and assisting academic writing for students who major in astronautics. To fulfill the objective, the author constructed a corpus by selecting 43 research articles published in *Journal of Guidance, Control and Dynamics (JGCD)* in 2020 and 2021. JGCD is one of the top journals in the field of astronautics and widely recognized by the industry. Its impact factor is 2.486 in 2021 and 2.809 for the last five years. Drawing on the most common textual metadiscourse markers in academic writing listed by Hyland [1], the author searched the self-built corpus by AntConc 3.5.9, and counted the standardized frequency results, so as to identify the most frequently used expressions. The theoretical significance of this study is to broaden the applying scale of Hyland's interpersonal model of interactive metadiscourse [1]. And the practical significance of this study lies in the beneficial guidance for academic writing in the field of astronautics.

## II. THEORETICAL BASIS

### A. Brief Introduction of Metadiscourse

"Metadiscourse" was first put forwarded by Zelling Harris in 1959. Since then, it has been depicted and defined from different perspectives and scopes. There exists a consensus now that metadiscourse binds writers, readers and discourse together, and enables readers a way to interpret the intended meaning in discourse. Metadiscourse is commonly believed to embody the connective and interpersonal relationships in discourse. Thus, by identifying these textual features, readers can examine the way the writer connects and organizes a text, and understand the value of specific discourse communities. Research of metadiscourse has been deeply influenced by Halliday's theory of three meta-functions of language. Scholars affected by this theory basically divided metadiscourse into two categories—textual category and interpersonal category. Based on systemic-functional grammar theory, this classification stressed that textual metadiscourse assists in constructing discourse while interpersonal metadiscourse functions to negotiate interaction and express stance [2].

From the statistics of Web of Science publications, before the 21st century there were not many metadiscourse academic papers published; after 2004, the related research began to increase steadily. Hyland [3] believes that this is related to the publication of two authoritative monographs: *Metadiscourse* [1]; *Metadiscourse in L1 and L2 English* [4]. In 2015, metadiscourse research increased exponentially and reached a peak in 2018, which happened to be the climax of "metadiscourse segmentation research", that is, researchers no longer only focused on all metadiscourse types of a certain text, but many researches focused on a certain type of metadiscourse emerged [5].

In recent years, domestic studies on metadiscourse have become a hotspot. Jiang [6] proposed the concept "metadiscoursive nouns" and emphasized its rhetorical function in both interactive and interactional dimensions, contributing a new analytical perspective on textual interaction in academic discourse. Based on a corpus of research articles across disciplines, this study found that "determiner + N" structure is most common, mainly used to set up anaphoric reference to prior discourse while expressing

authorial stance of defining research entities. Results also showed that metadiscursive nouns are more frequently used in the liberal arts than science and engineering domains in the rhetorical construction of social interaction and academic knowledge. Jiang Feng and Hyland [7] examined the change of interactional metadiscourse markers of academic prose among four disciplines between 1965 and 2015. They built a 2.2-million-word corpus of research articles from both hard and soft knowledge fields. Results showed that the overall trend of interactional metadiscourse is declining, with soft disciplines reducing but hard disciplines increasing. Lu Sitong [2] compared the textual features of research articles in different sub-domains (Second Language Acquisition and Discourse Analysis) within a discipline (Applied Linguistics) by drawing on Hyland's interpersonal model of interactive metadiscourse and its sub-types. Based on two self-constructed corpora by selecting research articles published in Applied Linguistics between 2014 and 2019, results showed that interactive metadiscourse was used significantly more in SLA than DA corpus. This study further testified the rhetorical and persuasive function of metadiscourse when constructing knowledge in nuanced specific research subjects, maintained the academic communication with disciplinary community.

From the brief review of recent studies of metadiscourse, conclusions can be drawn that metadiscourse is an efficient analytical framework for written discourse and has great potential in the future considering the lack of its clear theoretical clarification and relevant research tools for practical use. Interactional metadiscourse obtained more attention than interactive metadiscourse. There are relatively less studies that focus on the interactive metadiscourse. Research articles, abstracts, popular science articles, book reviews are investigated and research perspective are various, including cross-disciplinary approach, sub-domains-of-one-discipline approach, comparative approach, etc. Scholars prefer to construct their own small-sized corpus according to their research needs.

### B. Hyland's Interpersonal Model of Metadiscourse

To modify the drawbacks of previous classifications, Hyland [1] established the interpersonal model of metadiscourse in light of Thompson and Thetela's interactive and interactional resources.

Interactive resources are used to organize propositional information in ways that a projected target audience is likely to find coherent and convincing. They are clearly not simply text-organizing as their deployment depends on what the writer knows of his or her readers. They are a consequence of the writer's assessment of the reader's assumed comprehension capacities, understandings of related texts, and need for interpretive guidance, as well as the relationship between the writer and reader. There are five sub-categories [1]: transition markers, frame markers, endophoric markers, evidentials, and code glosses. The metadiscoursal devices enable parametric criteria for revealing how the writer organizes the text and negotiates author-readers interaction in writing practices. In this sense, Hyland's list of hundreds of potential metadiscourse items [1] provides a quantitative and operable method for researchers analyzing academic texts.

## III. RESEARCH DESIGN

### A. Research Questions

This study focuses on these two issues:

RQ1. What are the frequencies of each interactive metadiscourse used in the research articles of astronautics?

RQ2. What are typical forms of interactive metadiscourse used in the research articles of astronautics?

### B. Corpus Building

The building principle of this corpus is as follows. As a monthly journal, JGCD publishes 12 issues a year. For each month, articles in three columns, i.e., "Survey Papers", "Full-length Papers" and "Engineering Notes", were selected as the raw texts of the journal. Other columns, such as "Technical Comments", "Replies", "Corrections", "Article Updates", "Introductions" and "Editorials", were manually excluded. A total of 210 articles were published in 2020 and 190 in 2021.

The principle of sampling is as follows. Using stratified sampling, each year is divided into 12 layers according to the month. The sample ratio is set as 0.1, that is, the number of samples drawn per month (n) /the total number of journal articles per month (N) is 0.1 (that is, n/N=0.1). The number of samples in that layer is determined according to this proportion. And the automatic computer extraction program was used for random extraction. A total of 43 articles were sampled.

There are mainly three steps in the text preprocessing:

(1) Clean the garbled characters caused by mathematical formulas and pictures in the raw samples through Python;

(2) The tokenization of cleaned samples through CorpusWordParser;

(3) POS (Part-of-speech) annotation is completed online through CLAWS7 web page, using c7 tagset and horizontal output style.

### C. Research Procedure

According to Hyland's list of most common textual metadiscourse markers in academic writing [1], the author searched the corpus by AntConc3.5.9, a publicly free concordance. There are 173 interactive items on the list, which present explicit traces of textual interaction. While not exhaustive, the items provide a basis for examining intra-disciplinary variations and thus help to reduce the analytical complexity caused by the openness of metadiscourse [2] since "writers are able to add new items according to the needs of the context, and "the analyst may never recover all intended metadiscoursal meanings" [7]. In the end, the author calculated the frequency of each kind of interactive metadiscourse and listed the typical expressions of each kind.

## IV. RESULTS AND DISCUSSION

### A. Frequency of Interactive Metadiscourse in JGCD

In this self-built JGCD corpus, there are 19,601 word types and 200,814 word tokens. In Hyland's list [1], there are altogether 173 interactive items. The author identified 10,931 instances of the interactive metadiscourse, averaging about 54.43 cases per 1,000-word tokens. After searching item by item through AntConc 3.5.9, the frequency of each interactive metadiscourse are as follows (Table I).

Table 1. Frequency of interactive metadiscourse in the self-built JGCD corpus

| Classification of interactive metadiscourse | Frequencies of concordance hit (s) | Frequency/ 1,000-word tokens | Percentage of each classification |
|---|---|---|---|
| Frame markers | 1640 | 8.17 | 15.00% |
| Endophoric markers | 1527 | 7.60 | 13.97% |
| Evidentials | 53 | 0.26 | 0.48% |
| Code glosses | 538 | 2.68 | 4.92% |
| Transition markers | 7173 | 35.72 | 65.62% |

Notably, transition markers are most frequently used (35.72 per 1,000-word tokens), accounting for 65.62% of the total amount of interactive metadiscourse. The second most frequently used type is frame markers (8.17 per 1,000-word tokens), taking up 15.00%, followed by endophoric markers (7.60 per 1,000-word tokens, 13.97%). The frequency of evidentials is the lowest, i.e., 0.26 per 1,000-word tokens, accounting for 0.48%. According to statistics, for transition markers, the top three most frequently used expressions are "and (4,836 hits)" "also (334 hits)" "because (225 hits)". For frame markers, the top three most frequently occurred expressions are "first (222 hits)" "then (220 hits)" "(in) section X (182 hits)". As for endophoric markers, these are "Fig. X (796 hits)" "Figure X (190 hits)" "(In) Section X (182 hits)". In the scale of code glosses, the top three most frequently applied are "such as (128 hits)" "that is (104 hits)" "or X (71 hits)". For evidentials, there are only two items be identified: "according to X (52 hits)" and "cited (1 hit)".

It is worth mentioning that the overwhelmingly concordance hits of transition marker "and" to some extent undermines the attention to other markers. Among transition markers, besides the top three, the following expressions also obtain relatively high frequency: "however (199 hits)" "but (195 hits)" "thus (142 hits)" "therefore (106 hits)" "further (104 hits)" "although (99 hits)" "whereas (99 hits)".

### B. Discussion

#### 1) Typical forms of transition markers

Transition markers are mainly conjunctions and adverbial phrases which help readers interpret pragmatic connections between steps in an argument. Addition adds elements to an argument and potentially consists of items such as *and, furthermore, moreover, by the way,* etc. Comparison marks argument as either similar (*similarly, likewise, equally, in the same way, correspondingly,* etc.) or different (*in contrast, however, but, on the contrary, on the other hand,* etc.). Consequence relations either tell readers that a conclusion is being drawn or justified (*thus, therefore, consequently, in conclusion,* etc.) or that an argument is being countered (*admittedly, nevertheless, anyway, in any case, of course*) [1]. Transition markers also assist readers to understand connections between different steps in an argument [8]. Here are some typical forms of transition markers: Example 1 is the exemplification of transition markers performing additive relations. Example 2 represents the comparative ones.

Example 1: *Furthermore*, the performance of the aerocapture maneuver (in terms of impulsive V) is very sensitive to the bank-angle switching time).

Example 2: *Similarly* to the previous case the change in rotation regime is not sufficient to affect the surface environment 32%.

#### 2) Typical forms of frame markers

Frame markers signal text boundaries or elements of schematic text structure. Once again, care needs to be taken to identify features which order arguments in the text rather than events in time. Items included here function to sequence, label, predict and shift arguments, making the discourse clear to readers or listeners. Frame markers can therefore be used to sequence parts of the text or to internally order an argument, often acting as more explicit additive relations (*first, then, 1/2, a/b, at the same time, next*). They can explicitly label text stages (*to summarize, in sum, by way of introduction*). They announce discourse goals (*I argue here, my purpose is, the paper proposes, I hope to persuade, there are several reasons why*). And they can indicate topic shifts (*well, right, OK, now, let us return to*). Items in this category therefore provide framing information about elements of the discourse [1]. Thus, these devices are often primarily applied to order the sequence of the arguments or sections of a text as in Example 3. As shown in Example 4, additionally, they clearly label stages of the writer's unfolding presentation.

Example 3: *First*, as evident from Eq. (25), the control law includes a high-frequency switching action, which is not only detrimental to the integrity of the helicopter's mechanical control links and swash plates, but may also excite the high-frequency rotor modes.

Example 4: *In summary*, each aircraft will receive full information (state and action information) from other aircraft in the same sector, partial information (only state information) from other aircraft that are close to its own sector, and no information from all the remaining aircraft.

#### 3) Typical forms of endophoric markers

Endophoric markers are defined as "road signs of a text, pointing to and emphasizing different parts at different times" [9]. In other words, Endophoric markers are expressions which refer to other parts of the text. These make additional ideational material salient and therefore available to the reader in aiding the recovery of the writer's meanings, often facilitating comprehension and supporting arguments by referring to earlier material or anticipating something yet to come. By guiding readers through the discussion they help steer them to a preferred interpretation or reading of the discourse [1]. Endophoric references can be classified into linear and non-linear forms [2]. Linear references signify preceding or subsequent discourse and thus function as previews Example 5, reviews, or overviews of the unfolding text, whereas non-linear references as in Example 6 provide additional textual materials such as tables, figures, standalone examples or extracts

Example 5: For this reason, a discretization of the nonconservative force can be done by including these forces in the potential terms, as will be shown in "*typical forms of evidentials*".

Example 6: Figs. 8 and 9 present propagations performed with two different time steps.

### 4) Typical forms of evidentials

Evidentials guide the reader's interpretation and establish an authorial command of the subject. In some genres this may involve hearsay or attribution to a reliable source; in academic writing it refers to a community-based literature and provides important support for arguments [1]. Academic citations play a crucial role in demonstrating acknowledgment of usefulness and pertinence of prior quests for knowledge, and emphasize that research is embedded in a literature and that writers are linked into a wider disciplinary community [7]. Here present two typical forms of evidentials:

Example 7: This separation standard was chosen using the definition of well clear for UAS *according to* Cook and Brooks.

Example 8: A number of the *cited* examples for the different approaches to 3D path tracking are listed in Table 1, with a circle indicating that a certain property is in principle attainable using the method in question though it has not been demonstrated in the *cited* source.

### 5) Typical forms of code glosses

Code glosses supply additional information, by rephrasing, explaining or elaborating what has been said, to ensure the reader is able to recover the writer's intended meaning. They reflect the writer's predictions about the reader's knowledgebase and are introduced by phrases such as *this is called, in other words, that is, this can be defined as, for example,* etc. Alternatively, they are marked off by parentheses. To be more specific, code glosses are explanatory devices that construct coherent and reader-friendly texts, which mark writers' evaluation of readers' processing needs in digesting the academic texts [1]. Here are some typical forms of code glosses:

Example 9: The rotation minimizing frame is a special frame construction developed *specifically* to address the issues of frame discontinuities.

Example 10: Swarms may be indirectly adversarial conflicting with objectives simply due to not being under our control, as would be the case interacting with a natural swarm *such as* a bird flock, or with an outside team of vehicles.

## V. CONCLUSION

To sum up, this study adopts astronautics as a research case to investigate the textual features of research articles based on Hyland's interpersonal model of interactive metadiscourse [1], aiming to investigate the use of metadiscourse in research articles and assist academic writing for students who major in astronautics. A corpus consisting of 43 research articles published in *Journal of Guidance, Control and Dynamics* (JGCD) in 2020 and 2021 was built, containing 19,601-word types and 200,814 word tokens. And the author identified 10,931 instances of the interactive metadiscourse, averaging about 54.43 cases per 1,000-word tokens. Results show that transition markers are most frequently uesed (35.72 per 1,000-word tokens), accounting for 65.62% of the total amount of interactive metadiscourse, followed by frame markers (8.17 per 1,000-word tokens,

15.00%), endophoric markers (7.60 per 1,000-word tokens, 13.97%), and code glosses (2.68 per 1,000-word tokens, 4.92%). The frequency of evidentials is the lowest, i.e., 0.26 per 1,000-word tokens, accounting for 0.48%. This study also lists several frequently-used expressions of interactive metadiscourse for academic writing assistance. Students who are in need of verification of certain English expressions could also search in the self-built JGCD corpus.

After analyzing each classification of interactive metadiscourse, conclusions can be drawn as follows. First, transition markers play an important role in research articles in astronautics owing to the high requirement of logical expression. Second, the low frequency of evidentials indicates that research articles in astronautics are relatively more independent from others' opinions and concentrated on their own experiments and argumentations. This could be regarded as a sign of novelty to some extent. Third, the relatively high frequencies of frame markers and endophoric markers confirm that straightforward logic and clear structure together with transparent figures are preferred in research articles in astronautics.

As for the limitation of this study, in this study, there is no another authoritative assistant to double check the identification of interactive metadiscourse, implying the possible existence of mistakes during the identification process.

Suggestions for further research are as follow. To begin with, authoritative assistants should be included. Besides, the sample size of this self-built corpus could be enlarged for more universal conclusions about the textual features of research articles in astronautics.

## CONFLICT OF INTEREST

The author declares no conflict of interest.

## REFERENCES

[1] K. Hyland, *Metadiscourse: Exploring Interaction in Writing*, Bloomsbury Publishing, 2005.
[2] S. T. Lu, "The use of interactive metadiscourse in research articles across sub-domains of a discipline: A corpus-based case study of applied linguistics," M. S. thesis, School of Foreign Language Education, Jilin Univ., Jilin, China, 2022. (in Chinese)
[3] K. Hyland, "Metadiscourse: What is it and where is it going?" *Journal of Pragmatics*, vol. 113, pp. 16–29, May 2017.
[4] A. Ädel, *Metadiscourse in L1 and L2 English*, John Benjamins Publishing Company, 2006.
[5] S. M. Yu, "Bibliometric analysis of international metadiscourse studies within 60 years (1959–2019)," *Foreign Language Education*, vol. 42, no. 2, pp. 40–45, Mar. 2021. (in Chinese)
[6] F. Jiang, "Metadiscursive nouns: A new perspective on social interaction in academic discourse," *Journal of PLA University of Foreign Languages*, vol. 42, no. 2, pp. 63–159, Mar. 2019. (in Chinese)
[7] F. Jiang and K. Hyland, "Interactional metadiscourse: Argument and persuasion in changing academic contexts," *Foreign Language Education*, vol. 41, no. 2, pp. 23–28, Mar. 2020. (in Chinese)
[8] T. Dahl, "Textual metadiscourse in research articles: A marker of national culture or of academic discipline," *Journal of Pragmatics*, no. 10, pp. 1807–1825, 2004.
[9] A. Ädel and A. Mauranen, "Metadiscourse: Diverse and divided perspectives," *Nordic Journal of English Studies*, no. 2, pp. 1–11, 2010.