

<https://doi.org/10.17323/jle.2026.26937>

Comparative Analysis of Metadiscourse Markers in Moves of the Discussion Section of Chinese and English Research Articles across Disciplines

Jingfen Jin , Angkana Tongpoon Patanasorn 

Khon Kaen University, Khon Kaen, Thailand

ABSTRACT

Background. The Discussion section is where research findings are reported, interpreted, evaluated, connected to previous work, and turned into claims or implications. It is challenging for novice writers and EFL learners due to its complex rhetorical demands. Previous studies have focused on either rhetorical moves or metadiscourse markers (MDMs) in isolation, but much less is known about how particular markers help writers perform particular moves, and whether these patterns differ across languages and disciplines.

Purpose. This study examines how MDMs are distributed across rhetorical moves in the Discussion sections of Chinese and English research articles from soft and hard science. It seeks to identify disciplinary and linguistic differences in move-based marker use and explore how these patterns can inform genre-based academic writing instruction.

Method. The corpus included 200 Discussion sections from Scopus Q1 (English) and CNKI core (Chinese) journals. Texts were coded for rhetorical moves using an integrated move framework and for MDMs using Cao's taxonomy. Frequencies were normalized per 1,000 words. Chi square tests, Cramer's *V*, adjusted residuals, and qualitative examples were used to identify significant patterns.

Results. Disciplinary differences were move specific. In English articles, soft sciences used more attitude markers when drawing implications (15.17 vs 6.39), while hard sciences relied more on endophoric markers in background information and reporting results (3.88 vs 0.38 and 6.09 vs 1.69). In Chinese articles, soft sciences used more frame markers and boosters when reporting results (4.81 vs 1.53 and 8.05 vs 2.67), whereas hard sciences used more endophoric markers in reporting and commenting on results (6.76 vs 1.29 and 3.70 vs 0.64). The most striking cross-linguistic difference concerns self-mentions. English articles used them significantly more frequently than Chinese articles across several moves in both soft and hard sciences ($p < .001$, Cramer's $V=0.176-0.507$). An Argumentative Progression Pattern also emerged: writers shift from interactive to interactional markers as they develop interpretations, claims, evaluation and implications.

Conclusion. The findings demonstrate that MDM use in Discussion sections is shaped by discipline, language, and publication context. They also suggest that genre-based academic writing pedagogy should teach MDM as a move-sensitive resource, especially self-mentions, hedges, attitude markers, and endophoric markers.

KEYWORDS

rhetorical moves; metadiscourse markers; Discussion section; cross-linguistic variation; cross-disciplinary variation

Citation: Jin, J., & Patanasorn, A. T. (2026). Comparative analysis of metadiscourse markers in moves of the discussion section of Chinese and English research articles across disciplines. *Journal of Language and Education*, 12(1), 88-116. <https://doi.org/10.17323/jle.2026.26937>

Correspondence:
Angkana Tongpoon Patanasorn
angton@kku.ac.th

Received: May 05, 2025

Accepted: March 16, 2026

Published: March 31, 2026

INTRODUCTION

The Discussion section is an essential component of research articles (RAs), where authors interpret results, relate findings to previous studies, and artic-

ulate theoretical or practical implications (Amnuai, 2019; Mohammad, 2012; Nguyen & Pramoolsook, 2015; Yang & Allison, 2003; Zahra, Helen, & Ain, 2017). It is also widely regarded as one of the most rhetorically demanding sections of



an RA and plays a decisive role in manuscript acceptance for publication (Jin, 2018; Safnil, 2013). Given its importance, constructing an effective Discussion section still poses considerable challenges for many writers, particularly novice scholars and EFL writers such as Chinese early-career researchers. These challenges largely stem from the rhetorical complexity of this section and the need to align with disciplinary expectations and established academic conventions (Moreno & Swales, 2018; Soodmand, Doosti, & Movassagh, 2018).

One established approach to analyzing the rhetorical organization of Discussion sections is rhetorical move analysis, which identifies functionally distinct discourse segments that collectively realize the communicative purposes of a genre (Swales, 1990). Previous studies have shown that the Discussion section typically includes moves such as reporting results, interpreting findings, evaluating the present study, and drawing implications (Basturkmen, 2009, 2012; Hopkins & Dudley-Evans, 1988). While these studies provide valuable insights into the structural organization of Discussion sections, they also reveal considerable cross-disciplinary and cross-linguistic variation in terms of distribution, sequencing, and salience of rhetorical moves (Amnuai & Wannaruk, 2012; Peacock, 2002; Soodmand et al., 2018). Such findings suggest that rhetorical structure is shaped by disciplinary knowledge construction practices and academic writing traditions.

In addition to rhetorical structure, scholars have also examined the role of metadiscourse markers (MDMs), which are linguistic resources writers use to organize texts, express propositional stance, and engage readers in the texts (Hyland & Tse, 2004; Vande Kopple, 2002). Research on MDMs has demonstrated systematic cross-disciplinary and cross-linguistic variation. For example, soft sciences tend to employ more hedges, attitude markers, and engagement markers than hard sciences. This reflects differences in epistemological orientations and argumentative styles (Kadir, 2020; Kahkesh & Alipour, 2017; Tran & Duong, 2013). Cross-linguistic studies further indicate that English RAs often contain a wider range and higher frequency of MDMs (e.g., boosters and code glosses) than those written in other languages, such as Indonesian or Arabic (Sultan, 2011; Alharbi & Swales, 2011; Azizah, 2017; Faqih, 2022; Indarti, 2022; Kustiyasari et al., 2021; Nugrahani & Bram, 2020).

Although move analysis and MDM research have both contributed significantly to genre studies, these two research strands have largely developed independently. Most move-analytic studies focus primarily on discourse structure, whereas MDM research tends to examine linguistic features across entire sections or texts without reference to the rhetorical functions of specific discourse segments. This separation leaves an important dimension of academic discourse underexplored, namely, how MDMs are strategically deployed within particular rhetorical moves to achieve

specific communicative purposes. Correspondingly, little is known about how such move-embedded MDM patterns vary systematically across disciplines and languages.

A small number of studies have attempted to bridge this gap. For instance, Liu and Buckingham (2018) examined interactive and interactional MDMs in Applied Linguistics Discussion sections and demonstrated that MDMs serve as rhetorical tools for realizing move-specific communicative purposes. Although Ashofteh et al. (2020) also adopted an integrated analytical design, their study focused on the abstracts of research articles rather than Discussion sections. However, such studies remain limited in several respects. First, they typically focus on a single discipline, which makes it difficult to establish whether the observed patterns represent broader disciplinary tendencies. Second, cross-linguistic comparisons remain relatively scarce, especially between Chinese and English academic writing, despite the growing participation of Chinese scholars in international publication venues. Third, existing studies rarely explore how disciplinary knowledge and linguistic traditions jointly shape MDM distribution within specific rhetorical moves.

Addressing these limitations requires an analytical framework that integrates rhetorical structure with interpersonal meaning construction. Combining move analysis with MDM analysis offers such a framework. Rhetorical moves identify the communicative purposes writers pursue, while MDMs reveal how writers achieve these purposes and negotiate interpersonal relationships with readers linguistically. Therefore, examining the interplay between rhetorical moves and MDMs enables a more nuanced understanding of the rhetorical strategies underpinning academic persuasion. Conceptually, this approach integrates insights from Swales' (1990) genre theory and Hyland's (2005b) stance and engagement model. Linking rhetorical moves with MDMs also provides a more holistic account of how academic writers construct persuasive discourse across different rhetorical moves. In doing so, it responds to calls in genre and intercultural rhetoric research for analyses that capture the interplay between rhetorical structure and metadiscursive meaning across languages and disciplines (Hyland & Jiang, 2022; Li & Xu, 2022; Mur-Dueñas, 2011).

Against this backdrop, the present study examines how MDMs are distributed across rhetorical moves in the Discussion sections of Chinese and English research articles in the hard and soft sciences. The analysis focuses on four fields, namely Physics, Engineering, Education, and Linguistics, selected to represent a broad disciplinary spectrum in line with Biglan's classification. By bringing move analysis and MDM analysis into a single framework, the study offers a more detailed account of how writers use metadiscursive resources to perform move-specific communicative purposes. It also allows cross disciplinary and cross linguistic patterns to be examined within the same analytical design and provides a basis for tracing how MDMs contribute to the gradual de-

velopment of argumentation in Discussion sections. Accordingly, the study addresses the following research questions:

- RQ1:** What are the cross-disciplinary differences in MDM use within each rhetorical move in Chinese RAs?
- RQ2:** What are the cross-disciplinary differences in MDM use within each rhetorical move in English RAs?
- RQ3:** What are the cross-linguistic differences in MDM use within each rhetorical move in Chinese and English RAs of hard sciences?
- RQ4:** What are the cross-linguistic differences in MDM use within each rhetorical move in Chinese and English RAs of soft sciences?
- RQ5:** What argumentative progression pattern can be observed in the deployment of MDMs across rhetorical moves in the Discussion section of RAs?

LITERATURE REVIEW

Move Analysis of Discussion Sections

The rhetorical structure of the Discussion section has long been a focus in genre-based writing research, with move analysis serving as the primary methodological tool to reveal its communicative organization. Hopkins and Dudley-Evans (1988) proposed one of the earliest and most influential move frameworks, which is an eleven-move scheme developed from Biology dissertations and engineering conference papers. This framework comprises core moves such as Statement of Results, Expected/Unexpected Outcomes, and Explanation of Results, providing a detailed schema well suited to the rhetorical needs of hard sciences, where empirical results and direct explanations are usually the central parts. However, this framework follows a linear structure, lacks hierarchical distinctions between moves and steps, and has limited applicability to soft sciences that often rely on nuanced interpretation, layered argumentation, and contextualized discussion.

Subsequently, Yang and Allison's (2003) framework addressed some of these limitations by introducing a hierarchical movestep structure specifically derived from Applied Linguistics research articles. This seven-move framework includes units such as Commenting on Results and Drawing Implications, which align with the interpretive and dialogic nature of soft science discourse. It has been widely validated in soft science contexts (e.g., Hilmi, Toyyibah, & Afifi, 2021; Ulya, 2022; Yassaman & Elham, 2019). Despite its suitability for soft sciences, Yang and Allison's (2003) framework fails to capture key rhetorical features prevalent in hard sciences, such as explicit emphasis on Unexpected Findings and formal Claims. These are moves central to disciplinary communication in fields such as Chemistry and Biology.

Existing studies have confirmed that Discussion move frameworks are largely disciplinespecific. Hopkins and Dudley-Ev-

ans' (1988) scheme better accommodates hard science rhetoric, whereas Yang and Allison's (2003) framework is more applicable to soft sciences. Nevertheless, there remains a lack of flexible, integrated move frameworks suitable for cross-disciplinary comparison, as neither fully captures rhetorical diversity across hard and soft science domains. Building on these seminal works, the present study incorporates core components from both frameworks to develop a synthesized eight-move framework. This integrated framework mitigates the limitation of disciplinary boundedness by integrating hard science-oriented moves (e.g., expected/unexpected results) with soft science-aligned hierarchical movestep organization. Accordingly, it supports a robust analytical design adaptable to systematic cross-disciplinary and cross-linguistic comparisons.

Metadiscourse Frameworks and Classifications

Metadiscourse markers (MDMs) are significant rhetorical resources for constructing persuasive and interpersonal academic discourse, and their functional classification has evolved substantially to capture their pragmatic complexity. Vande Kopple's (1985) seminal taxonomy laid the theoretical foundation by distinguishing textual (organizational) from interpersonal (stance-taking) markers. This binary classification aligns with Halliday's (1973) systemic functional metafunctions (ideational, interpersonal, textual) and offers a clear conceptual starting point, yet it lacks fine-grained subcategory differentiation.

Crismore et al. (1993) later refined Vande Kopple's (1985) taxonomy by restructuring subcategories and clarifying the functional scope of each type. Though more descriptively comprehensive, its internal complexity and overlapping categories constrained practical application in empirical coding, as researchers frequently encountered ambiguity when classifying multifunctional MDMs. A major breakthrough came with Hyland's (2005a, 2019) influential taxonomy, which divides MDMs into interactive markers (transitions, frame markers, endophoric markers, code glosses, evidentials) and interactional markers (hedges, boosters, attitude markers, self-mentions, engagement markers). This framework balances clarity and comprehensiveness, resolving many limitations of earlier classifications and gaining widespread adoption across academic genres. Even so, it does not fully elaborate subtle MDM sub-types recognized in Hyland's earlier works, such as nuanced distinctions among evidential and hedge variants.

Cao (2014) extended Hyland's (2005a, 2019) taxonomy by refining sub-types and elaborating functional descriptions grounded in Hyland's earlier works (1996a, 1996b, 2005b, 2007). This extended framework distinguishes linear versus non-linear endophoric markers, integral versus non-integral evidentials, clarifies multifunctional self-mention use, expands hedge functional distinctions, and provides a standardized MDM wordlist for consistent coding. These

refinements mitigate practical drawbacks inherent in previous taxonomies. Scholarly consensus recognizes Hyland's (2005a, 2019) classification as the dominant foundational framework for MDM research, while Cao's (2014) extension enhances coding reliability and analytical granularity. Furthermore, this framework has been empirically validated across multiple contexts (Cao & Hu, 2014; Hu & Cao, 2015; Hyland & Jiang, 2022). Therefore, the present study adopts Cao's (2014) taxonomy as the analytical framework for precise MDM coding.

Cross-Disciplinary Variation in Academic Writing

Cross-disciplinary variation in academic writing, particularly within Discussion sections, has long been a core in genre-based research. Many empirical studies have consistently identified divergent rhetorical organization and linguistic practice shaped by disciplinary epistemologies. Soft science fields (e.g., Applied Linguistics, Sociology) prioritize interpretive, dialogic discourse, as knowledge construction hinges on nuanced argumentation, contextualized interpretation, and engagement with multiple perspectives. In contrast, hard sciences (e.g., Chemistry, Biology) emphasize empirical demonstration, objective reporting, and direct result-conclusion alignment, reflecting positivist epistemological underpinnings.

Empirical research has indicated distinct cross-disciplinary patterns in Discussion rhetorical moves. Soft science Discussions frequently feature Commenting on Results (M4), Evaluating the Study (M7), and Drawing Implications (M8), which facilitate interpretive and dialogic engagement (Gao & Pramoolsook, 2022; Soodmand et al., 2018; Zahra et al., 2017). These moves closely co-occur with both interactional and interactive MDMs, including hedges for epistemic caution, self-mentions for authorial presence, and frame markers for argument structuring (Kahkesh & Alipour, 2017; Tran & Duong, 2013). Conversely, hard science Discussions often open with Background Information (M1) to situate findings within existing scholarship, while prioritizing evidential markers to anchor claims to empirical data and endophoric markers to reference internal textual elements (Amnuai & Wannaruk, 2012).

Collectively, prior studies demonstrate that disciplinary epistemologies shape both rhetorical move structures and MDM deployment in Discussion sections, with persistent disparities between hard and soft science traditions. What remains underexplored is how such disciplinary differences intersect with MDM use within specific rhetorical moves, that is, whether the functional linkage between moves and MDMs varies systematically across disciplinary domains. The present study builds on relative studies by focusing disciplinary variation (hard vs. soft sciences) and adopting a move-specific lens to examine MDM deployment, to address the limitation of isolating move analysis from MDM investigation.

Cross-Linguistic Variation in Metadiscourse Markers

Cross-linguistic variation in MDM use has attracted growing scholarly attention, especially in English as a Foreign Language (EFL) context, where researchers examine how linguistic and cultural conventions shape writers' rhetorical choices. A dominant research trajectory compares MDM deployment between international English research articles (by native or near-native English scholars) and EFL-authored English RAs across diverse linguistic backgrounds and reveals cross-linguistic patterns.

Studies in EFL contexts (e.g., Algeria, Indonesia) report lower frequencies of interactional MDMs, such as self-mentions, engagement markers, and boosters, in EFL writing relative to international English RAs (Kadir, 2020; Azizah, 2017; Faqih, 2022). Such findings indicate that EFL writers tend to adopt more tentative, less overtly authoritative rhetorical stances, potentially influenced by cultural norms of modesty or limited familiarity with Anglophone academic discourse conventions. However, existing cross-linguistic research predominantly focuses on overall MDM frequency rather than move-specific distribution. Consequently, how cross-linguistic disparities manifest within individual rhetorical moves of the Discussion section remains unclear.

Prior work confirms cross-linguistic divergence in MDM practice, particularly the underuse of interactional MDMs in EFL writing. Nonetheless, it remains unclear how these cross-linguistic patterns fluctuate across rhetorical moves and interact with disciplinary fields. Specifically, whether cross-linguistic MDM variation holds consistently across both hard and soft sciences. Drawing on this literature, the present study frames cross-linguistic difference as a key analytical variable by comparing MDM use in Chinese (EFL) and international English RAs to address the oversight of move-MDM interdependence in conventional cross-linguistic MDM research.

Integrative Approaches Combining Moves and Metadiscourse

While prior research has examined rhetorical moves and MDMs separately, an integrative approach (linking move analysis to MDMs classification) offers a more nuanced understanding of academic discourse. Move analysis identifies the "what" of communicative practice (i.e., the rhetorical functions writers perform), while MDMs analysis reveals the "how" (i.e., how these functions are realized persuasively and interpersonally) (Hyland, 2005a; Devitt, 2015). This integration bridges the gap between structural (move) and linguistic (MDMs) analyses and enables researchers to explore how rhetorical functions are linguistically enacted.

Based on the communicative functions of rhetorical moves and the functions of MDM types, potential associations be-

tween specific rhetorical moves and MDM types are outlined below. For example, Background Information (M1) and Evaluating the Study (M7) may use self-mentions to establish authorial presence or acknowledge limitations; Reporting Results (M2) may employ boosters to assert findings and frame markers to organize multiple outcomes; Commenting on Results (M4) and Making Claims (M5) may use hedges to express caution, evidentials to link arguments to prior research, and boosters to reinforce claims; and Drawing Implications (M8) may deploy hedges to soften recommendations and frame markers to signal organization. A conceptual diagram (see Figure 1) illustrates these anticipated associations, with the inner ring representing the eight rhetorical moves from the combined framework and the outer ring presenting the MDM types most typical to each move. This diagram is informed by the communicative purposes of moves and functional characteristics of MDMs (Cao, 2014).

Obviously, integrating move analysis and MDM classification provides a more comprehensive understanding of academic discourse than either approach alone, as it links structural and linguistic levels of analysis. What remains unresolved is how these move-MDM associations vary across disciplines and languages, a gap that limits our understanding of cross-cultural and cross-disciplinary academic communication. Thus, we adopt the integrative approach (moves + MDMs) as our core conceptual framework by using the combined eight-move model (synthesized from Hopkins & Dudley-Evans, 1988 and Yang & Allison, 2003) and Cao’s (2014) MDM taxonomy to examine move-specific MDM use, addressing the limitation of prior studies that failed to integrate these two perspectives.

While prior studies have explored rhetorical moves, MDM use, cross-disciplinary variation and cross-linguistic differences, little research has examined how MDMs function within specific rhetorical moves across both languages and disciplines. This gap is critical, as it limits our understand-

ing of how disciplinary epistemologies and linguistic contexts jointly shape the rhetorical and linguistic practices of academic Discussion sections (a central part of effective knowledge construction and communication). The integrated framework (combined move model + Cao’s MDMs taxonomy) is necessary to address this gap. This integrated design enables us to examine cross-disciplinary (hard vs. soft sciences) and cross-linguistic (Chinese vs. English) variation in MDM use within specific rhetorical moves and to address the identified research gap. This study is also helpful to advance genre-based academic writing research and provide practical implications for EFL instruction and cross-cultural academic communication.

METHOD

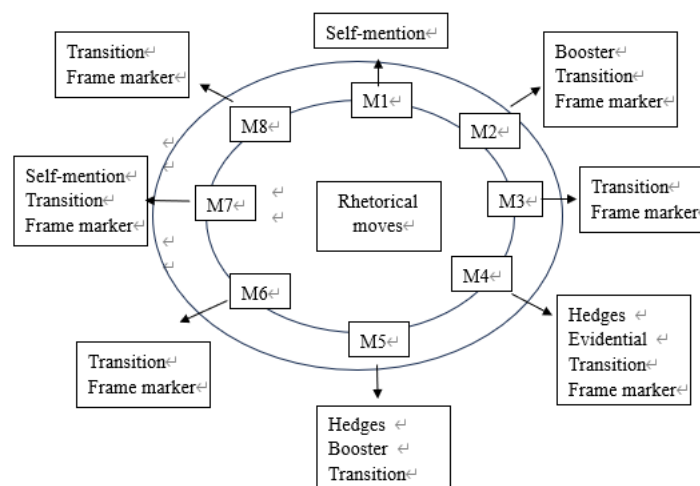
Corpus

This study employed two self-built corpora: the English Research Article Discussion Corpus (EN-RAD) and the Chinese Research Article Discussion Corpus (CH-RAD). All texts were collected from publicly accessible peer-reviewed academic journals. The dataset consists exclusively of published research articles and does not involve human participants, personal data or confidential information. As such, the study presents no ethical risks to individuals or institutions. The use of these materials complies with established ethical standards for corpus-based and genre analysis research, where publicly available texts are analyzed for scholarly purposes. All sources were used strictly for academic research only.

Each corpus consists of 100 research article Discussion sections, which yields a total dataset of 200 texts. Within each corpus, 50 articles were drawn from hard sciences (Physics and Engineering) and 50 from soft sciences (Education and Linguistics). These four disciplines were selected to represent Biglan’s (1973) disciplinary classification, namely hard-pure

Figure 1

Hypothesized preferred MDMs across rhetorical moves in the Discussion section



(Physics), hard-applied (Engineering), soft-applied (Education) and soft-pure (Linguistics). From each discipline, 25 articles were randomly selected. Compared with many previous move-analytic studies that typically analyze 10 to 20 articles per discipline (e.g., Amnuai, 2019; Basturkmen, 2012; Holmes, 1997; Liu & Buckingham, 2018; Yang & Allison, 2003), the present dataset offers broader disciplinary coverage and is sufficiently large to support a manually annotated corpus study.

To enhance corpus representativeness and comparability, all articles were sampled from well-established peer-reviewed journals within their respective disciplines. English-language articles were selected from Q1 journals indexed in Scopus, as this quartile comprises top-ranked international journals (e.g., Ahmed, 2023). Chinese-language articles were chosen from core journals indexed in the China National Knowledge Infrastructure (CNKI), which represent the leading scholarly outlets within the Chinese academic publication system. The selected journals are widely recognized within their respective disciplinary communities and follow rigorous peer-review procedures. Detailed information on the selected journals is provided in Appendix A.

Eventually, the EN-RAD corpus contains 139,721 running words, while the CH-RAD corpus comprises 182,111 running words. Across disciplines, the number of running words ranges from 27,686 to 40,205 in EN-RAD and from 33,025 to 54,894 in CH-RAD (see Table 1). This distribution ensures relatively balanced disciplinary representation and provides a robust dataset for move identification and MDM analysis.

The present study acknowledges an inherent design limitation in corpus sampling. The Chinese corpus is drawn from domestic CNKI core journals, whereas the English corpus comprises articles from international Scopus Q1 journals. This mismatch introduces an unavoidable confounding factor. Consequently, cross-linguistic comparisons in this study cannot be interpreted as reflecting purely linguistic differences. Instead, they represent the combined influence of intertwined dimensions: language, publication context (domestic versus international venues) and academic conventions (local Chinese scholarly norms versus global English academic discourse practices). This is recognized as a methodological design limitation rather than a mere interpretive caveat, and should be taken into consideration when interpreting the findings and generalizing the results.

Article Selection Criteria

For article selection, four criteria were followed. First, only articles published between 2022 and 2023 were included to capture recent rhetorical practices in disciplinary academic writing. As diachronic changes may exert an impact on rhetorical conventions (Biber, Ulla, & Thomas, 2007; Jalilifar, Hayati, & Namdari, 2012), the time span was limited to reduce such confounding effects.

Second, author background was carefully controlled. Chinese articles were required to be written exclusively by Chinese scholars and published in Chinese-medium journals, while English articles were drawn from international English-medium journals and could be authored by scholars of any nationality. This distinction reflects the different communicative contexts represented in the two corpora. The Chinese dataset aims to capture local Chinese academic discourse practices, whereas the English dataset represents international English-language scholarly communication, which functions as the dominant medium of global academic publication. Author nationality and linguistic background were verified using available author information, such as biographical notes, institutional affiliations, and academic profiles (Gao & Pramoolsook, 2022).

Third, the publication context of the journals was considered to ensure comparability between the two datasets. Chinese articles were retrieved from the CNKI database and selected from reputable peer-reviewed academic journals indexed in major Chinese academic journal lists. English articles were obtained from internationally recognized Q1 journals indexed in Scopus, which are widely used indicators of international scholarly visibility. In both cases, the journals were well-established within their respective disciplines and followed rigorous peer-review procedures. This approach ensured that the selected texts represented comparable levels of academic quality, disciplinary legitimacy, and publication standards, regardless of the different linguistic contexts.

Fourth, to maintain genre consistency, all selected articles had to be empirical research studies (primarily quantitative in nature) and follow the conventional IMRD (Introduction–Method–Results–Discussion) structure (Soodmand et al., 2018). In addition, only research articles with a clearly labeled Discussion section were included in the corpus. Ar-

Table 1
Running Words in the Corpus

	CH-RAD				EN-RAD			
	SA	SP	HP	HA	SA	SP	HP	HA
Word count	54,894	54,558	33,025	39,634	34,363	37,467	27,686	40,205
Soft/hard	109,452		72,659		71,830		67,891	
Total	182,111				139,721			

ticles with merged sections (e.g., Results and Discussion or Discussion and Conclusion) were excluded to maintain clear rhetorical boundaries and facilitate more reliable move identification (Ahmadi, 2022; Loi, Evans, Lim, & Akkakoson, 2016). These procedures ensured that the two corpora were comparable in terms of disciplinary scope, publication quality, and genre conventions.

Procedure

First, the PDF versions of eligible articles were downloaded, and the Discussion sections were manually extracted into Microsoft Word. A second round of screening was then conducted to verify the boundaries of the Discussion sections and ensure consistency across articles. During this process, minor typographical errors and irregular spacing were corrected. However, tables, figures, and reference lists were removed to retain only the running text for analysis.

Subsequently, the Discussion sections were manually annotated for rhetorical moves using the combined eight-move framework described in Appendix B. The framework was developed by synthesizing previous frameworks for research article Discussions and was operationalized through a set of explicit coding guidelines that specify the communicative purpose and linguistic indicators of each move.

Prior to formal coding, a pilot annotation was conducted on a subset of articles to familiarize the coders with the framework and refine the operational definitions of each move. Then, move annotation was carried out independently by two trained coders with experience in academic discourse analysis. During the coding process, segments were identified according to their rhetorical function rather than strictly by sentence boundaries, as moves can be realized at the level of sentences, clauses, and phrases (Swales, 1990). When ambiguous segments occurred (e.g., sentences simultaneously referring to results and providing interpretation), coders first consulted the operational definitions of the moves and examined the broader textual context. When disagreements arose, they were discussed jointly until consensus was reached.

To illustrate ambiguous cases and the resolution of systematic disagreements, a representative example is provided. A sentence such as “The observed results partially support our hypothesis, which is consistent with prior research but also differs in several key aspects” was initially coded differently by the two coders: one coder classified it as M4 (Commenting on Results) due to its interpretive nature, while the other coded it as M5 (Making Claims) because it advances a tentative conclusion. To resolve this disagreement systematically, the coders first revisited the operational definitions in the coding scheme: M4 focuses on interpreting results in relation to prior studies, while M5 emphasizes asserting novel claims. The two coders then analyzed the broader textual context and noted that the sentence primarily elaborates

on how the results align with and diverge from previous studies (a core function of M4) rather than making a new claim. Through this systematic review of coding criteria and contextual analysis, the coders reached a consensus to code the segment as M4. This approach, relying on operational definitions, contextual analysis, and joint discussion, was consistently applied to resolve all systematic disagreements during the coding process. In addition, several borderline cases were reviewed again after the initial round of coding to ensure the consistent application of the framework.

After the completion of move coding, the texts were reorganized according to moves. Specifically, segments belonging to the same move were compiled into separate files to facilitate subsequent MDM analysis. This procedure resulted in eight move-based sub-corpora within each disciplinary cluster, which allows for the examination of MDMs within specific rhetorical functions. MDMs were then identified and coded following Cao’s (2014) classification scheme and a detailed coding manual. The identification of MDMs also underwent a pilot-to-formal annotation process. Ten general MDM categories (see detailed information in Appendix C) were examined thoroughly. The identification of MDMs was guided by both lexical lists and contextual interpretation, as certain expressions may function differently depending on their rhetorical context.

This two-stage procedure (first move annotation and then MDM coding within each move) enabled a systematic examination of how MDMs are distributed across rhetorical moves in the Discussion sections. For the entire procedure, Microsoft Word was used for text cleaning, segmentation, and preliminary coding, while Microsoft Excel was used for frequency calculations, data organization, and graphical visualization. Statistical analyses (chi-square tests, which were used to examine cross-linguistic and cross-disciplinary differences) were conducted using SPSS (Version 27).

To illustrate the annotation procedure, Excerpt 1 demonstrates how rhetorical moves were identified within the Discussion section.

Excerpt 1: This study examined how fourth-grade students experienced cognitive load over the course of a reading comprehension test administered in three different test formats [**M1: Background Information**]. This study found no differences in experienced cognitive load between reading comprehension tests administered on paper, on the computer, or in a computer-adaptive format. Students experienced more cognitive load at the end of each test than in the middle [**M2: Reporting Results**]. This intuitive effect offers further support for the working memory resource depletion effect (citation anonymized for review purposes), as it can be reasoned that working memory resources deplete over the course of a test, meaning that more mental effort is required to complete test items [**M4S3: Explaining Results**]. The increase in cognitive load did not differ between test media (computer versus pencil-and-paper), but when comparing the adaptive test format with the fixed-item test format, there was limited evidence for the tendency that cognitive load may increase more in an adaptive test, though the

effect was very small and not consistent across models [M2: Reporting Results].

After move coding was completed, all segments belonging to the same move were compiled into separate files for MDM analysis. For instance, all M2 segments (Reporting Results) were grouped together for the identification of MDMs. Excerpt 2 illustrates the annotation of MDMs within an M2 segment.

Excerpt 2: This study found **7BOO** no differences in experienced cognitive load between reading comprehension tests administered on paper, on the computer, or in a computer-adaptive format. Students experienced more cognitive load at the end of each test than in the middle [M2]. The increase in cognitive load did not differ between test media (computer versus pencil-and-paper) **5REF**, but **1COM** when comparing the adaptive test format with the fixed-item test format, there was limited evidence for the tendency that cognitive load may **6HED** increase more in an adaptive test, though **1COM** the effect was very small and not consistent across models [M2].

Reliability Measurement

To ensure the reliability of the coding process, both inter-rater and intra-rater agreement were assessed. The primary researcher and a second coder (a PhD student in Applied Linguistics, a native Chinese speaker and fluent English user) first completed a rigorous training phase, in which they jointly coded two sample articles and resolved discrepancies before working on the main dataset. Following training, a random 30% of the corpus was independently double-coded, consistent with the procedures outlined by Mackey and Gass (2022). Inter-rater reliability was reported in terms of raw agreement rates and Cohen's Kappa values. For move coding, the agreement rate reached 90.46% for the CH-RAD and 92.38% for the EN-RAD ($\kappa = 0.80$, $\kappa = 0.83$, $p < .001$). For MDM coding, the rates were 86.05% for the CH-RAD and 88.56% for the EN-RAD ($\kappa = 0.75$, $\kappa = 0.77$, $p < .001$). Intra-rater reliability was also strong, with agreement rates of 95.89% (CH-RAD) and 97.24% (EN-RAD) for move coding ($\kappa = 0.90$, $\kappa = 0.92$, $p < .001$), and 97.78% (CH-RAD) and 98.45% (EN-RAD) for MDM coding ($\kappa = 0.93$, $\kappa = 0.95$, $p < .001$). These values indicate satisfactory inter-rater and intra-rater agreement, confirming a high degree of coding consistency across raters and over time (Cohen, 1988; Landis & Koch, 1977).

Data Analysis

The analysis examined ten general categories of MDMs across eight rhetorical moves in the Discussion sections of research articles by focusing on two dimensions, namely, disciplinary domain (soft vs. hard sciences) and language (Chinese vs. English). RQ1 and RQ2 were answered by exploring the cross-disciplinary differences within each language corpus, while RQ3 and RQ4 were tackled through cross-linguistic comparisons within each disciplinary field. Based on the data for RQ1 and RQ2, an argumentative pro-

gression pattern (RQ5) was summarized to illustrate MDM deployment across the sequential development of rhetorical moves.

MDM raw frequencies were normalized per 1,000 words for each move and presented in tables and figures to visualize cross-group patterns. Chi-square tests were conducted using SPSS (Version 27) to examine group differences. Given the presence of zero-frequency cells, a small constant (0.5) was added to stabilize the distribution and meet chi-square assumptions, a standard correction procedure in corpus-based frequency analysis. As several expected cell frequencies were lower than five, the Monte Carlo simulation method was employed for all 2 (disciplinary/language groups) \times 10 (MDM types) contingency tables across the eight rhetorical moves. Effect sizes were reported using Cramer's *V*, which were interpreted according to Cohen's (1988) conventional thresholds (.10 = small, .30 = medium, .50 = large). The interpretation of Cramer's *V* depends on degrees of freedom (*df*) that are calculated as $df = \min(r-1, c-1)$, where *r* and *c* denote the number of rows and columns in the contingency table. In the present 2 \times 10 tables, $df = \min(2-1, 10-1) = 1$. It renders that Cohen's standard thresholds are applicable. Additionally, Post hoc comparisons were performed using adjusted standardized residuals (*ASRs*). For the 2 \times 10 tables, a Bonferroni correction was applied ($\alpha = .05/20 = .0025$). Accordingly, an absolute *ASR* value ≥ 3.09 was taken to indicate that a given cell contributed significantly to the overall chisquare result.

RESULTS

Chinese Soft vs. Hard Sciences

Figure 2 presents the normalized frequencies (per 1,000 words) of MDMs in Chinese soft and hard sciences. Overall, clear disciplinary differences can be observed. Soft science articles tend to employ more frame markers (4.81 vs. 1.53) and boosters (8.05 vs. 2.67) in M2 (Reporting Results), as well as slightly more evidentials (7.79 vs. 6.02) and self-mentions (0.50 vs. 0.48) in M4 (Commenting on Results). In contrast, hard science articles favor endophoric markers (6.76 vs. 1.29 in M2; 3.70 vs. 0.64 in M4) and engagement markers (4.59 vs. 1.67 in M2; 5.39 vs. 2.78 in M4), and also use hedges more frequently in M8 (Drawing Implications) (10.56 vs. 3.35). The normalized frequency differences between soft and hard sciences reflect their distinct rhetorical needs. soft sciences rely more on frame markers, boosters, and self-mentions to convey research findings, while hard sciences tend to use endophoric markers, engagement markers, and hedges to ensure the rigor and objectivity of research results.

It is noteworthy that although self-mentions were detected in the Chinese corpus, their frequency remained extremely low (fewer than one occurrence per 1,000 words). This sug-

gests that Chinese research articles generally maintain an impersonal rhetorical style.

Chi-square tests revealed significant associations between discipline and MDM distribution in M2 ($\chi^2(9, N = 1021) = 216.094, p < .001, \text{Cramer's } V = 0.460$), M4 ($\chi^2(9, N = 2812) = 189.212, p < .001, \text{Cramer's } V = 0.259$), and M8 ($\chi^2(9, N = 432) = 28.236, p < .001, \text{Cramer's } V = 0.256$) (see Table 2). These values indicate an approaching large effect association between discipline and MDM distribution in M2 and approaching medium effects in M4 and M8. Post hoc analyses (Bonferroni-adjusted) confirmed that these differences were primarily driven by several specific categories. Soft science articles used frame markers and boosters more frequently than expected in M1 (Background Information) and evidentials and self-mentions in M2, whereas hard science articles showed a significantly higher use of endophoric markers and engagement markers in M2 and M4, as well as hedges in M8 ($|ASR| > 3.09$). Detailed observed frequencies, expected frequencies, and *ASR* values are presented in Appendix D1. The significant associations observed in chi-square tests and post hoc analyses suggest that disciplinary characteristics directly influence MDM deployment. The approaching large effect association in M2 and approaching medium effects in M4 and M8 further confirm that discipline-specific requirements shape the choice of MDMs, which is consistent with the functional differences between soft and hard sciences.

The following examples further illustrate these tendencies. Soft science writers often use frame markers to organize their presentation of findings, as in:

Example 1: 其次, ICT干预对听力理解的干预效果最低, 且对听力理解干预的泛化效果处于极弱水平。(Secondly, the ICT intervention had the lowest effect on listening comprehension, and its transfer effect was extremely weak.)

They also use evidentials to relate their findings to previous studies:

Example 2: 这一研究发现与De Koning (2010)等人以往的研究结果相一致。(This finding is consistent with the results of previous studies by de Koning et al. (2010))

Differently, hard science articles frequently guide readers to data through endophoric markers:

Example 3: 从表2可以看出, 无论是12%Cr系列还是13%Cr系列材料, 随着Mo含量的提高两次回火后其中的逆变奥氏体含量变化不大。(From Table 2, it can be seen that the content of reverted austenite changes very little with the increase in Mo content.)

Moreover, hedges often appear in concluding moves when discussing possible improvements:

Example 4: 此外, 网络模型有一定的改进空间, 将来或许可以通过采集更少的点来进行预测分析。(In addition, the network model still has room for improvement, and it may be pos-

sible to conduct predictive analysis with fewer data points in the future.)

English Soft vs. Hard Sciences

Figure 3 illustrates the normalized frequencies of MDMs in English soft and hard science articles. Compared with soft sciences, hard science articles use more endophoric markers in M1 (3.88 vs. 0.38) and M2 (6.09 vs. 1.69). This indicates a stronger reliance on references to figures and tables when contextualizing the research and presenting results, reflecting that hard sciences emphasize data-driven presentation and guiding readers toward objective evidence. Hard science writers also employ more attitude markers in M4 (3.59 vs. 1.76). This suggests that they tend to adopt a clearer evaluative stance when commenting on results. By comparison, soft science articles deploy more attitude markers in M8 (15.17 vs. 6.39). This indicates soft science writers are more inclined to express evaluative attitudes when concluding the Discussion section.

As shown in Table 3, chi-square analyses revealed significant associations between disciplinary domain and MDM distribution in M1 ($\chi^2(9, N = 716) = 17.388, p < .05, \text{Cramer's } V = 0.156$), M2 ($\chi^2(9, N = 1815) = 58.493, p < .001, \text{Cramer's } V = 0.180$), M4 ($\chi^2(9, N = 3316) = 34.616, p < .001, \text{Cramer's } V = 0.102$) and M8 ($\chi^2(9, N = 1251) = 28.821, p < .001, \text{Cramer's } V = 0.152$). The corresponding effect sizes were generally small. The significant chi-square results confirm that disciplinary domain also shapes MDM distribution in English research articles, mirroring the pattern observed in the Chinese corpus. Nevertheless, the generally small effect sizes indicate that disciplinary differences in English RAs are less pronounced than those found in the Chinese corpus. This implies that the influence of discipline on MDM use may vary across linguistic contexts.

Post hoc tests confirmed that these differences were mainly attributable to the MDM categories outlined above. Hard science articles used endophoric markers significantly more often in M1 ($ASR = 3.8$) and M2 ($ASR = 6.9$), while attitude markers occurred more frequently in M4 ($ASR = 3.3$). Conversely, soft science articles exhibited higher-than-expected use of attitude markers in M8 ($ASR = 4$). Appendix D2 provides detailed information on observed frequencies, expected frequencies, and *ASR* values. The post hoc results further validate the specific MDM categories driving disciplinary differences in English RAs. These findings align with the functional roles of these MDMs and reinforce the relationship between disciplinary rhetorical expectations and MDM deployment in English RAs.

For instance, endophoric markers in hard science texts often guide readers to visual data:

Figure 2

Normalized frequency of MDMs in Chinese Soft vs. Hard Science

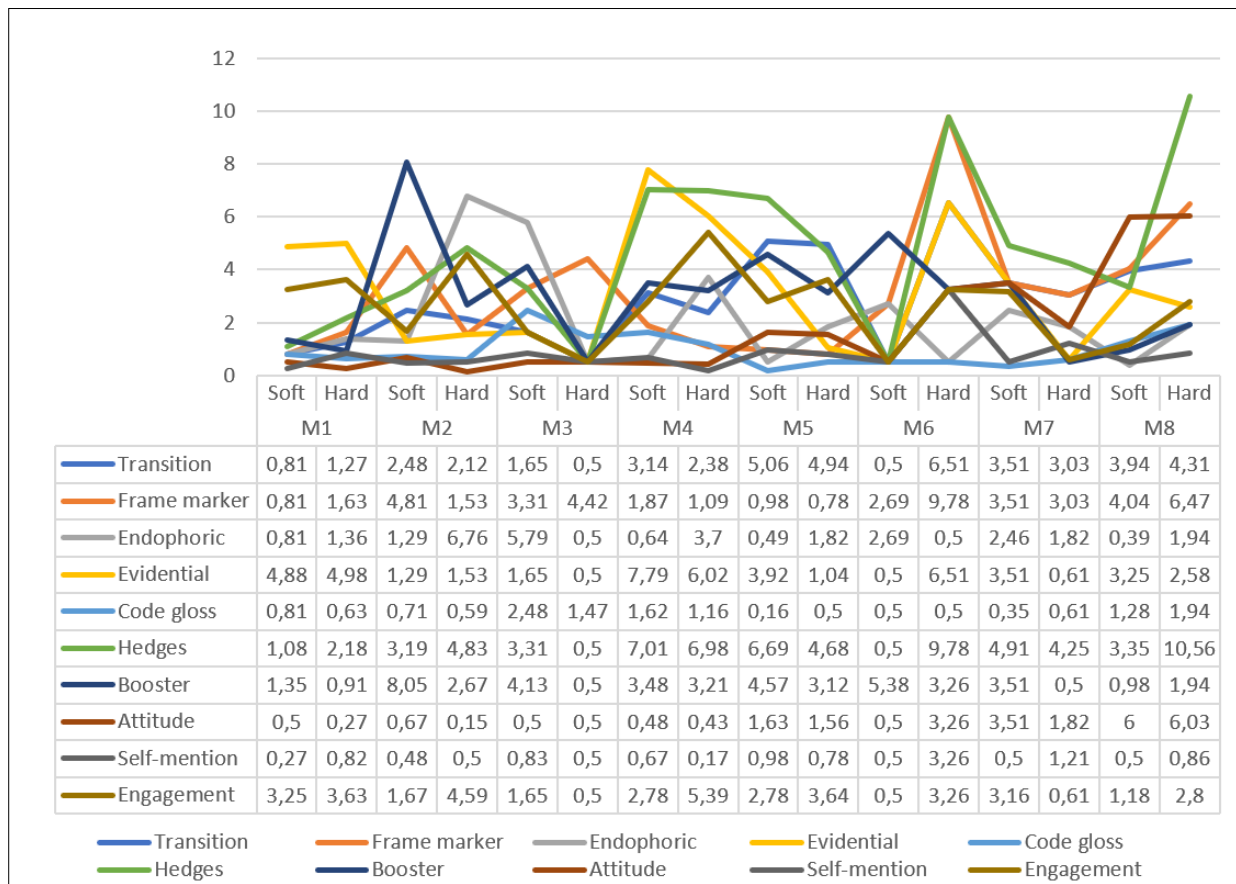


Table 2

Statistical Results of Chinese Soft vs. Hard Science

Move	χ^2	Monte Carlo p	Cramer's V	Significant categories (ASR)
M1	4.922	= .852	0.141	—
M2	216.094	< .001	0.460	Frame marker, Booster (S _↑ , H _↓), Endophoric, Engagement (H _↑ , S _↓)
M3	3.373	= .976	0.280	—
M4	189.212	< .001	0.259	Evidential, Self-mention (S _↑), Endophoric, Engagement (H _↑ , S _↓)
M5	13.572	= .136	0.231	—
M6	2.148	= .997	0.282	—
M7	9.314	= .420	0.290	—
M8	28.236	< .001	0.256	Hedges (H _↑ , S _↓)

Note. ↑ means significantly higher than expected; ↓ means significantly lower than expected.

Example 5: The probability of heat transfer was sampled stochastically from the temperature versus probability of ignition curve in Fig. 6b.

Attitude markers also appear when evaluating the significance of findings:

Example 6: It is striking, but probably not fortuitous.

Chinese vs. English Soft Sciences

Figure 4 demonstrates the normalized frequencies (per 1,000 words) of MDMs in Chinese and English soft science articles. A striking cross-linguistic difference concerns the use of self-mentions. English articles consistently employ self-mentions across several moves, including M1 (13.38 vs. 0.27), M2 (9.04 vs. 0.48), M4 (4.54 vs. 0.67), M7 (28.55 vs.

Figure 3

Normalized frequency of MDMs in English Soft vs. Hard Science

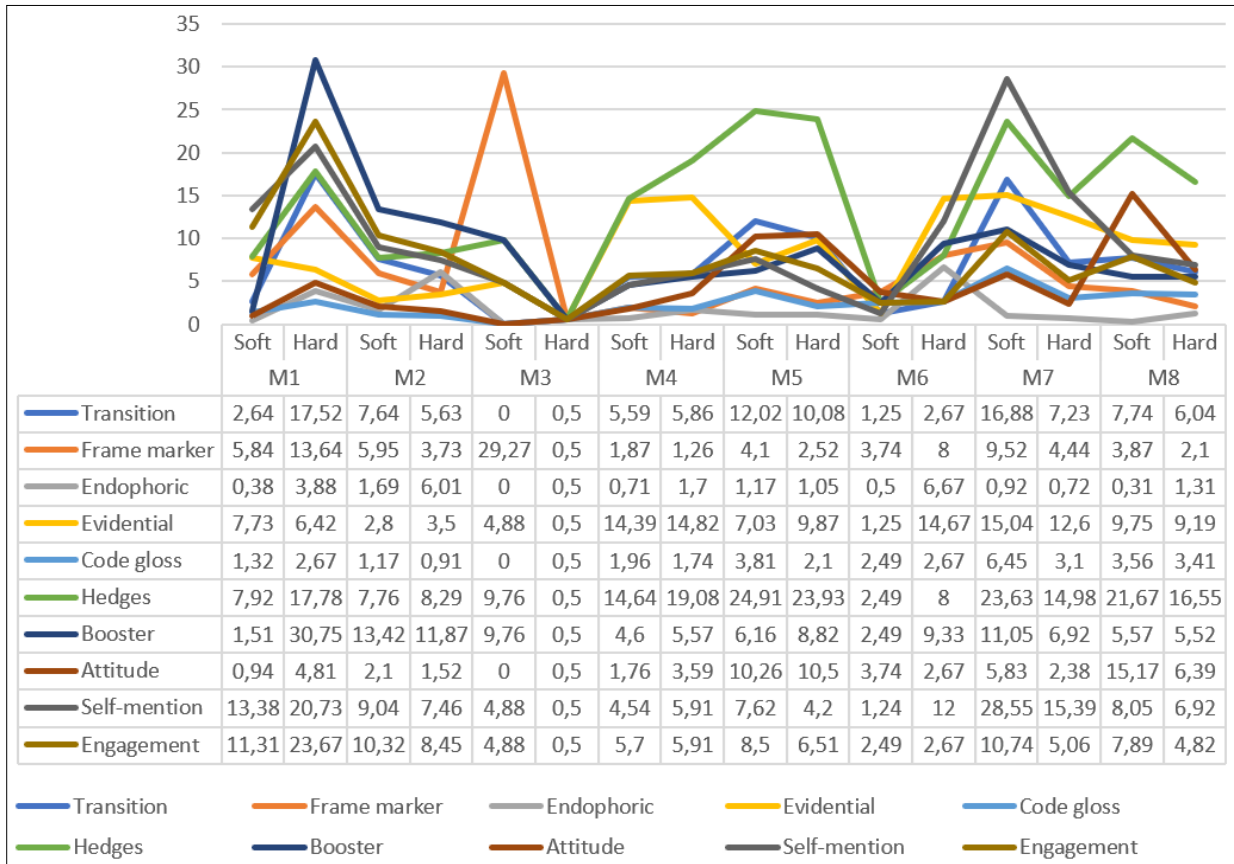


Table 3

Statistical Results of English Soft vs. Hard Science

Move	χ^2	Monte Carlo p	Cramer's V	Significant categories (ASR)
M1	17.388	< .050	0.156	Endophoric (H \uparrow , S \downarrow)
M2	58.493	< .001	0.180	Endophoric (H \uparrow , S \downarrow)
M3	2.598	= .990	0.310	—
M4	34.616	< .001	0.102	Attitude (H \uparrow , S \downarrow)
M5	12.307	= .196	0.135	—
M6	9.419	= .419	0.367	—
M7	11.93	= .216	0.103	—
M8	28.821	< .001	0.152	Attitude (S \uparrow , H \downarrow)

Note. \uparrow means significantly higher than expected; \downarrow means significantly lower than expected.

0.50), and M8 (8.05 vs. 1.18). The higher use of self-mentions across multiple moves in English soft science articles compared with their Chinese counterparts suggests a far stronger authorial presence in English academic writing. This cross-linguistic difference reflects distinct rhetorical norms for constructing authorial identity in soft science research across linguistic contexts.

Chi-square analyses confirmed significant associations between language and MDM distribution in M1 ($\chi^2 (9, N = 334) = 37.654, p < .001, \text{Cramer's } V = 0.336$), M2 ($\chi^2 (9, N = 1,578) = 132.28, p < .001, \text{Cramer's } V = 0.290$), M4 ($\chi^2 (9, N = 3,854) = 118.768, p < .001, \text{Cramer's } V = 0.176$), M5 ($\chi^2 (9, N = 459) = 29.012, p < .001, \text{Cramer's } V = 0.251$), M7 ($\chi^2 (9, N = 501) = 49.855, p < .001, \text{Cramer's } V = 0.315$) and M8 ($\chi^2 (9, N =$

789) = 82.851, $p < .001$, Cramer's $V = 0.324$) (see Table 4). Effect sizes were approaching medium for M2, medium for M7 and M8, and small for the remaining moves. The significant chi-square results demonstrate that language is a key factor shaping MDM distribution in soft science Discussion sections. The medium effect sizes for M2, M7, and M8 indicate that cross-linguistic differences in MDM use are most prominent within these rhetorical moves, while small effects for other moves seem to reveal a consistent yet subtler linguistic influence. Collectively, these findings confirm that language plays a vital role in shaping MDM deployment in soft science Discussion sections.

Post hoc analyses further showed that English articles used self-mentions significantly more frequently than expected across these moves ($|ASR| > 3.09$). Detailed observed frequencies, expected frequencies, and ASR values are presented in Appendix D3. The post hoc results confirm that self-mentions are one of the primary MDM categories driving cross-linguistic differences in soft science articles.

Some apparent discrepancies between normalized frequency patterns and post hoc results were observed for markers such as evidentials, boosters, and frame markers. For instance, although boosters yielded higher normalized frequencies in English articles in M2 (13.42 vs. 8.05) and M4 (4.60 vs. 3.48), post hoc tests revealed that Chinese articles used these markers significantly more frequently than expected. This discrepancy arises because normalized frequencies reflect absolute usage rates, whereas post hoc tests identify relative overrepresentation within the contingency table. These findings suggest that Chinese writers tend to cluster certain MDMs within specific rhetorical moves, whereas English writers maintain more consistent MDM deployment throughout the Discussion section.

This contrast is also reflected in the examples. English authors frequently refer explicitly to their research activities and interpretations:

Example 7: Regarding **our** second aim, we address students' behavior in relation to the use of instructional components and resources of the programs.

Differently, Chinese writers tend to maintain a more impersonal tone when describing research procedures:

Example 8: 本次行动研究开展了两轮基于移动平台的视听续说教学…… (This action research conducted two rounds of mobile platform-based audiovisual retelling instruction...)

Chinese vs. English Hard Sciences

Figure 5 displays the normalized frequencies (per 1,000 words) of MDMs in Chinese and English hard science articles. As in soft sciences, self-mentions are considerably more frequent in English articles across several moves: M1 (20.7 vs. 0.82), M2 (7.46 vs. 0.50), M4 (5.91 vs. 0.17), M7 (15.4 vs. 1.21), and M8 (6.91 vs. 0.86). The higher frequency of self-men-

tions across multiple rhetorical moves in English hard science articles indicates a consistently strong authorial presence in English hard science writing. This pattern aligns with the findings for soft sciences. Collectively, the frequent use of self-mentions in English articles represents a cross-linguistic feature that transcends disciplinary boundaries. This reflects divergent rhetorical norms for constructing authorial identity in English and Chinese academic discourse.

Chi-square analyses revealed significant associations between language and MDM distribution in M1 ($\chi^2(9, N = 630) = 40.001, p < .001$, Cramer's $V = 0.252$), M2 ($\chi^2(9, N = 1,258) = 143.173, p < .001$, Cramer's $V = 0.337$), M4 ($\chi^2(9, N = 2,274) = 584.311, p < .001$, Cramer's $V = 0.507$), M5 ($\chi^2(9, N = 466) = 31.246, p < .001$, Cramer's $V = 0.259$), M7 ($\chi^2(9, N = 734) = 36.325, p < .001$, Cramer's $V = 0.222$) and M8 ($\chi^2(9, N = 894) = 71.551, p < .001$, Cramer's $V = 0.283$) (see Table 5)). The significant chi-square results suggest that language appears to shape MDM distribution in hard science Discussion sections. The medium effect sizes in M2 and large effect sizes in M4 suggest that cross-linguistic differences in MDM use are most pronounced within these core rhetorical moves. Small-to-medium strength associations for the remaining moves further indicate that language may exert a consistent influence across other stages of the Discussion. Overall, these findings demonstrate that language systematically affects patterns of MDM deployment in hard science Discussions.

Post hoc tests showed that English articles used self-mentions significantly more frequently than expected in M1, M2, M4, and M8 ($|ASR| > 3.09$). Detailed statistical information is provided in Appendix D4. The post hoc results confirm that self-mentions are the primary MDM category driving cross-linguistic differences in hard science articles.

Another notable cross-linguistic difference concerns endophoric markers, which appeared somewhat more frequently in Chinese hard science articles across M2, M4, M5, and M7. Nevertheless, their overall normalized frequencies remained comparatively low. This may suggest that these markers were not extensively deployed in either corpus.

The following examples illustrate the differences in authorial presence. English authors frequently refer explicitly to their research process:

Example 9: Our study involves testing the effect of two mediating variables on the relationship between the dependent and the independent variables.

Comparatively, Chinese articles typically adopt a more neutral and impersonal style:

Example 10: 在体模实验中, 利用光声断层扫描成像系统从外部扫描体模…… (In the phantom experiment, a photoacoustic tomography imaging system was used to scan the phantom externally...)

Figure 4

Normalized frequency of MDMs in Chinese vs. English soft Science

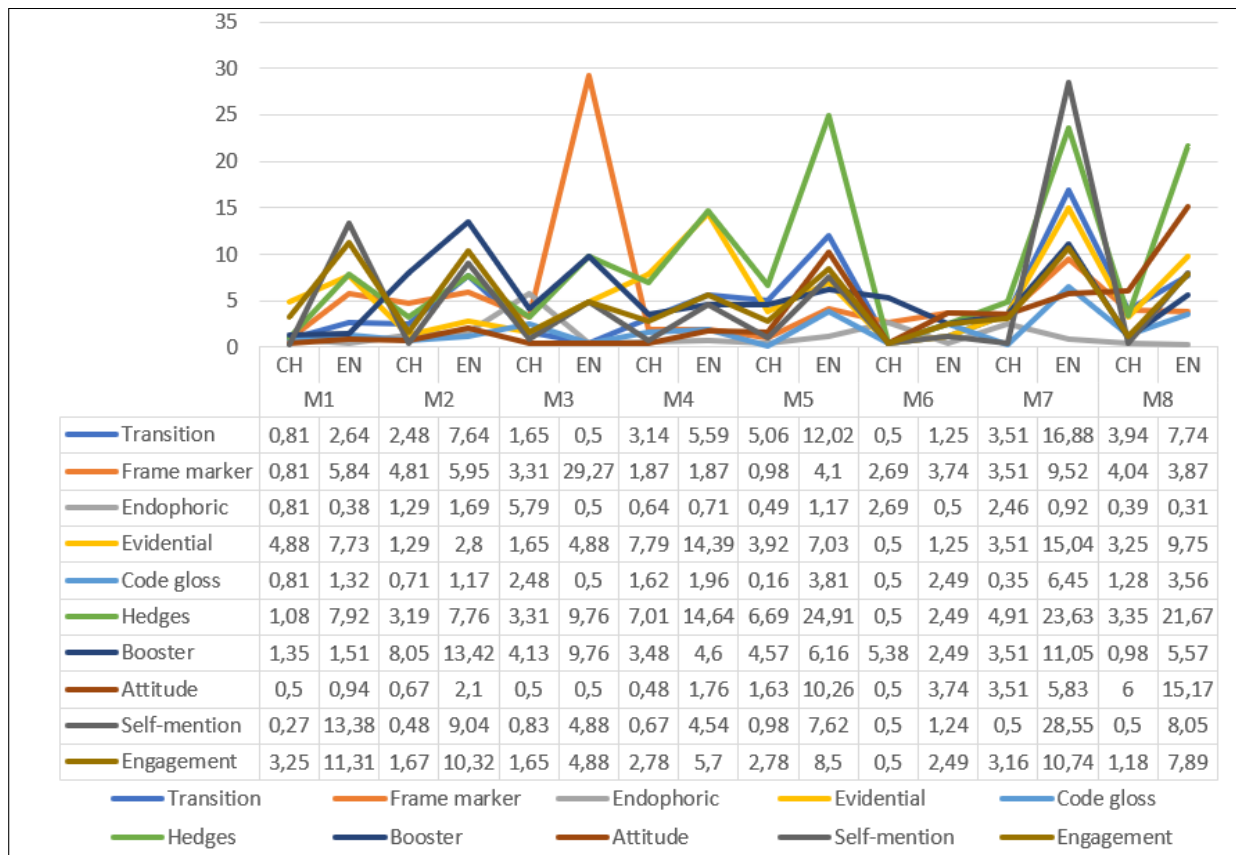


Table 4

Statistical Results of Chinese vs. English Soft Science

Move	χ^2	Monte Carlo p	Cramer's V	Significant categories (ASR)
M1	37.654	< .001	0.336	Evidential (CH \uparrow , EN \downarrow), Self-mention (EN \uparrow , CH \downarrow)
M2	132.28	< .001	0.290	Frame marker, Booster (CH \uparrow , EN \downarrow), Self-mention, Engagement (EN \uparrow , CH \downarrow)
M3	0.829	= .995	0.169	—
M4	118.768	< .001	0.176	Frame marker, Booster (CH \uparrow , EN \downarrow), Self-mention (EN \uparrow , CH \downarrow)
M5	29.012	< .001	0.251	Booster (CH \uparrow , EN \downarrow)
M6	1.391	= .995	0.219	—
M7	49.855	< .001	0.315	Endophoric (CH \uparrow , EN), Self-mention (EN \uparrow , CH \downarrow)
M8	82.851	< .001	0.324	Frame marker (CH \uparrow , EN \downarrow), Hedge, Self-mention (EN \uparrow , CH \downarrow)

Note. \uparrow means significantly higher than expected; \downarrow means significantly lower than expected.

Argumentative Progression Pattern

The previous analyses mainly focused on cross-linguistic and cross-disciplinary comparisons. The present analysis further reveals a tentative, data-driven descriptive pattern, which is named the Argumentative Progression Pattern (for MDM deployment across rhetorical moves in the Discussion section). As the discussion develops from contextualizing results to drawing implications, different MDM types gain

dominance in distinct moves. This pattern directly reflects how preferred MDMs interact with the communicative functions of individual rhetorical moves to advance logical argumentation throughout the Discussion section.

Concretely, in the opening moves of M1 (Background Information) and M2 (Reporting Results), writers employ interactive markers such as transitions, frame markers, and ev-

Figure 5

Normalized frequency of MDMs in Chinese vs. English hard Science

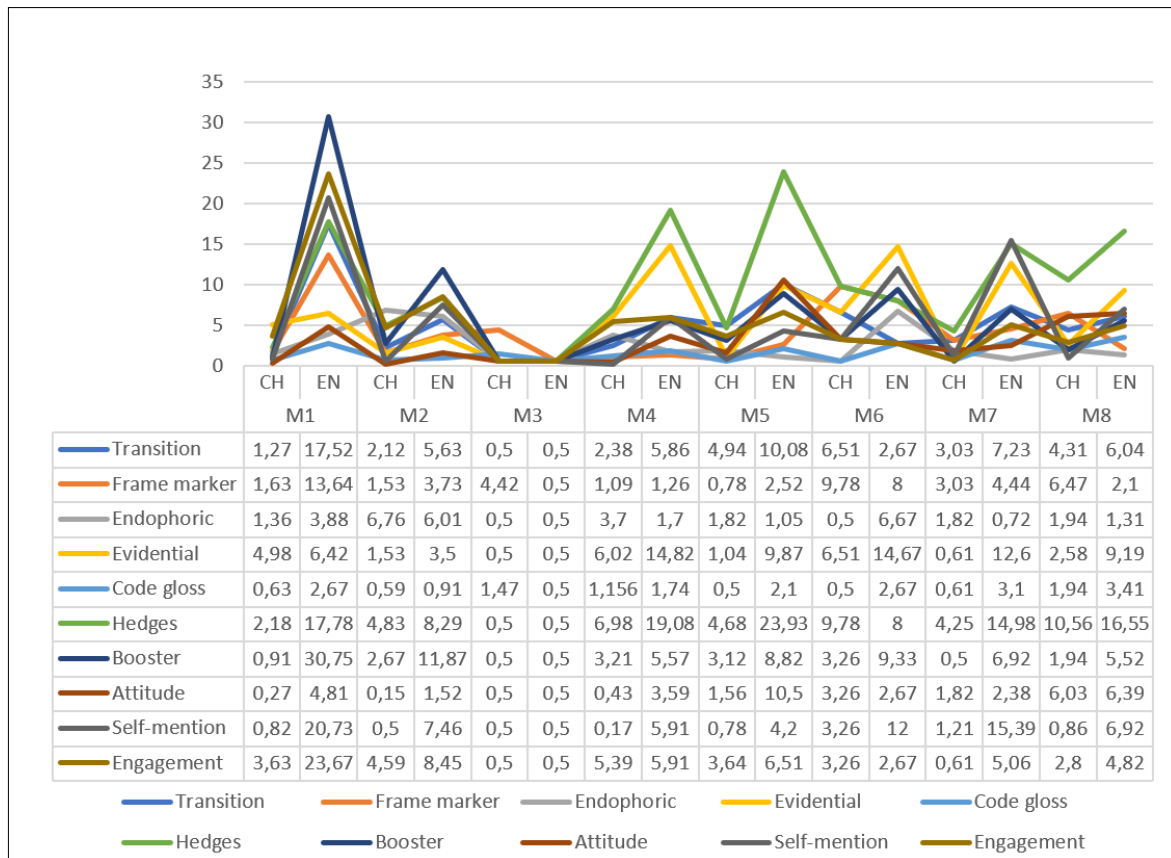


Table 5

Statistical Results of Chinese vs. English Hard Science

Move	χ^2	Monte Carlo p	Cramer's V	Significant categories (ASR)
M1	40.001	< .001	0.252	Evidential (CH \uparrow , EN \downarrow), Self-mention (EN \uparrow , CH \downarrow)
M2	143.173	< .001	0.337	Endophoric (CH \uparrow , EN), Booster, Self-mention (EN \uparrow , CH \downarrow)
M3	2,726	= .995	0.258	—
M4	584.311	< .001	0.507	Transition, Endophoric, Evidential, Engagement (CH \uparrow , EN \downarrow), Frame, code gloss, hedge, attitude, self-mention (EN \uparrow , CH \downarrow)
M5	31.246	< .001	0.259	Endophoric (CH \uparrow , EN \downarrow)
M6	5.195	= .864	0.276	—
M7	36.325	< .001	0.222	Endophoric (CH \uparrow , EN \downarrow)
M8	71.551	< .001	0.283	Frame marker (CH \uparrow , EN \downarrow), Self-mention (EN \uparrow , CH \downarrow)

Note. \uparrow means significantly higher than expected; \downarrow means significantly lower than expected.

identials to organize discourse and contextualize empirical findings. In M3 (Summarizing Results), frame markers are also frequently deployed to signal the transition from result presentation to interpretive discussion and to guide readers through textual structure. As the discussion proceeds to M4 (Commenting on Results), hedges and evidentials become more prevalent. This tendency reflects a need to offer cautious interpretations while establishing links with prior

scholarship. In M5 (Making Claims), interactional markers including boosters and attitude markers appear more frequently to articulate explicit authorial stance. When summarizing the study (M6), frame markers, again play a vital role in structuring discourse and highlighting the transition toward evaluative commentary. Subsequently, self-mentions and boosters are strategically used to foreground authorial positioning and reinforce the significance of key findings in

M7 (Evaluating the Study). Finally, in the concluding move M8 (Drawing Implications), hedges and engagement markers prevail when writers present practical and theoretical implications and invite readers to engage in broader disciplinary discussions. Overall, these regularities demonstrate that MDMs function as essential rhetorical resources underpinning the sequential development of argumentation in Discussion sections.

These empirical observations are summarized into the tentative Argumentative Progression Pattern (see Figure 6). This pattern visualizes how MDMs are strategically deployed to support staged argumentative development across rhetorical moves, with the inner ring representing the eight rhetorical moves in the Discussion section and the outer ring presenting the MDM types most typical to each move (namely, rhetorical moves are arranged in the sequential argumentative order of the Discussion section; dominant MDM types for each move are presented concisely to illustrate the rhetorical progression). Compared with the hypothesized MDM-move associations initially proposed in Figure 1, the present empirical results yield a data-informed refinement of this diagram. Importantly, this tentative pattern is different from conventional accounts of Discussion structure. Most prior scholars predominantly focus on the linear sequencing of rhetorical moves (e.g., Hopkins & Dudley-Evans, 1988; Yang & Allison, 2003) and treat individual moves as isolated functional units. By contrast, the present pattern reveals a dynamic, sequential interplay between MDM types

and move-specific communicative purposes. This pattern demonstrates that the dominant MDMs within each rhetorical move are not randomly distributed but closely aligned with shifting argumentative goals (from contextualization to implication-drawing). Accordingly, this tentative descriptive pattern offers a fresh insight into how academic writers construct persuasive discourse at the intersection of rhetorical moves and MDMs and advances relative literature by transcending mere descriptions of rhetorical move sequencing.

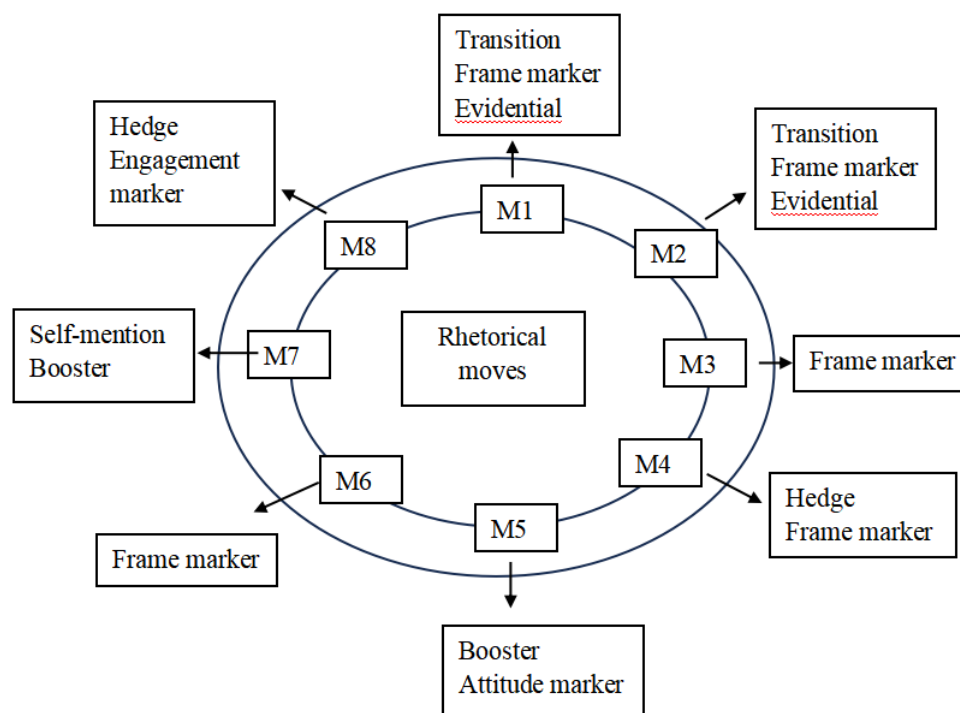
DISCUSSION

Cross-Disciplinary Comparison

The findings reveal clear disciplinary variations in the distribution of metadiscourse markers (MDMs) across rhetorical moves. By examining the interaction between rhetorical moves and MDMs, the results suggest that disciplinary writing practices influence not only the overall frequency of MDMs but also the ways in which these resources are strategically positioned within specific rhetorical moves in the Discussion section. Rather than merely documenting disciplinary disparities in MDM use, the present study further situates these observed patterns within existing genre and MDM studies, which extends earlier research that has examined rhetorical move structures and MDM distributions as two independent dimensions (e.g., Kahkesh & Alipour, 2017; Soodmand et al., 2018).

Figure 6

A tentative argumentative progression for MDM deployment in the Discussion section



We found writers in Chinese soft science texts employed more frame markers and boosters in M2 (Reporting Results) and more evidentials in M4 (Commenting on Results) than their hard science counterparts. Rather than simply reflecting stylistic preferences, this pattern may relate to the interpretive orientation of soft disciplines, where authors are typically expected to contextualize findings, engage with theoretical perspectives and position their results within ongoing scholarly debates. Therefore, frame markers can help organize extended arguments, boosters emphasize the relevance or significance of findings, and evidentials situate results within existing literature. These functions support the interpretive and dialogic nature of knowledge construction in soft sciences. Claims in this field are frequently developed through argumentative discussion rather than solely through empirical demonstration. This observed tendency largely confirms the disciplinary epistemological accounts proposed by Becher & Trowler (2001), who argued that soft science knowledge is interpretive, contested and discourse-dependent. Furthermore, the denser deployment of frame markers, boosters and evidentials identified in the present study extends the empirical observations of Hyland (2005a) and Peacock (2002) by demonstrating that such soft-science MDM preferences are not only overall textual tendencies but also move-specific rhetorical choices concentrated in result-reporting and result-commenting stages (Kahkesh & Alipour, 2017).

Differently, endophoric markers occurred more frequently in Chinese hard science texts in M2 and M4 and in English hard science texts in M1 (Background Information) and M2. These markers typically direct readers to figures, tables or other sections of the article. In so doing, the narrative interpretation can be integrated with the empirical evidence. Such patterns are consistent with the epistemological orientation of hard sciences, where arguments are closely tied to experimental data, visual representations and methodological procedures (Hyland, 2004). The higher concentration of endophoric markers across early rhetorical moves in hard science Discussions directly supports Hyland's (2004) claim that hard science writing prioritizes empirical anchorage and transparent reference to in-text data. In this context, endophoric references function as an organizational device that supports data-driven reasoning and facilitates readers' navigation through complex empirical information, providing further empirical validation for conventional assumptions about positivist knowledge construction in hard science academic writing.

Another disciplinary difference concerns a higher frequency of engagement markers and hedges in Chinese hard science articles. Engagement markers in M2 and M4 often restate or clarify information. This potentially helps readers navigate technically dense descriptions. Such a strategy suggests that even in highly technical contexts, writers employ engagement markers to guide readers through complex explanations and maintain textual coherence. Hedges, particularly

in M8 (Drawing Implications), mitigate claims and signal caution when discussing limitations or future research directions. These patterns indicate that even disciplines often characterized as objective or impersonal still rely on stance and engagement markers to manage the presentation of knowledge claims, although such texts are usually composed in more restrained forms. This finding partially aligns with Hyland and Tse (2004) and extends their arguments by showing that hedging and reader engagement in hard sciences are not randomly distributed but are strategically concentrated in interpretive and implication-generating rhetorical moves (Hyland, 2001).

Disciplinary differences of English articles also emerged in the use of attitude markers. English soft-science writers employed them more frequently in M8 (Drawing Implications), whereas English hard-science writers tended to use them in M4 (Commenting on Results). This distribution suggests that evaluative commentary in soft sciences often appears when authors discuss broader implications, recommendations or theoretical significance, whereas in hard sciences evaluation tends to be integrated more directly into the interpretation of results. This move-specific distribution of attitude markers strongly corroborates the cross-disciplinary findings reported by Peacock (2002) and Amnuai & Wannaruk (2012), and further refines their general observations by pinpointing the exact rhetorical moves where disciplinary evaluative preferences systematically manifest.

These findings support the view that disciplinary genres function as distinct communicative systems shaped by different knowledge-construction practices (Swales, 1990; Bhatta, 1993). Discussion sections in soft disciplines often serve as spaces for interpretation, contextualization and theoretical positioning, whereas those in the hard sciences remain more closely tied to empirical data and emphasize precision, methodological transparency and cautious qualification of claims (Basturkmen, 2012).

From a pedagogical perspective, these patterns highlight the importance of teaching novice writers how MDMs function strategically within disciplinary discourse, rather than treating them as a purely stylistic feature. In particular, writing instruction should help students recognize how different rhetorical moves in the Discussion section require different types of MDMs. For example, frame markers and boosters may be particularly useful for organizing and emphasizing interpretive arguments in soft science discussions, whereas endophoric markers and hedges may play a more prominent role in guiding readers through data interpretation and cautiously presenting claims in hard-science contexts. Moreover, embedding corpus-based examples from published research articles into classroom instruction can help learners observe how experienced writers deploy MDMs in specific rhetorical moves. Such activities can raise students' awareness of how stance, engagement and textual organization are shaped by disciplinary conventions. By analyzing

ing authentic disciplinary texts, novice writers can develop greater rhetorical flexibility and learn to adapt their MDM strategies to the expectations of their academic communities.

Cross-Linguistic Comparison

The cross-linguistic comparison reveals systematic differences in MDM use between Chinese and English research articles, particularly in self-mention deployment. This study contrasts Chinese-language articles by domestic Chinese scholars with English-language articles from international Q1 journals with culturally and linguistically diverse contributing authors. Observed disparities stem not merely from linguistic differences but also from the combined influences of language norms, genre conventions, academic training, and publication contexts.

English articles employ self-mentions significantly more frequently across M1, M2, M4, M7, and M8 than Chinese articles. English writers project stronger authorial presence throughout the Discussion section, whereas Chinese scholars avoid explicit first-person reference and favor impersonal framing to maintain a neutral rhetorical tone. The present findings further confirm the cross-linguistic tendency identified by Mu et al. (2015) regarding English writers' greater reliance on self-mention and salient authorial identity. This study further extends their work by locating such disparities within specific rhetorical moves, a dimension overlooked in broad cross-linguistic comparisons of authorial stance alone.

Notable discrepancies emerge between normalized MDM frequencies and post hoc pairwise comparisons for evidentials, boosters, and frame markers. Although English articles yield higher normalized booster frequencies in result-reporting and result-commenting moves, post hoc analyses reveal significantly greater-than-expected booster use in Chinese texts within these same moves. Overall English MDM frequency is generally higher, while Chinese writing exhibits concentrated clustering of specific markers in individual rhetorical moves. Normalized frequencies capture global usage trends, whereas post hoc tests expose context-specific overrepresentation within move-bound discourse functions. These dual statistical perspectives refine existing cross-linguistic MDM research, which has largely depended solely on raw or normalized frequency counts. By integrating both global distribution and move-specific clustering, the present method offers a more granular account of how writers strategically allocate MDMs across the argumentative stages of the Discussion section. That is, English writers distribute MDMs evenly across rhetorical phases, whereas Chinese writers deploy individual marker categories more selectively to fulfil localized rhetorical functions.

English articles exhibit a higher overall density of MDMs relative to Chinese counterparts. This pattern aligns with established cross-linguistic scholarship on academic meta-discourse (Alharbi & Swales, 2011; Azizah, 2017; Faqih, 2022; Indarti, 2022; Kustyasari et al., 2021). The current results confirm the consensus in prior literature that English academic writing employs a broader range and higher frequency of interactive and interactional MDMs than non-English academic traditions. Cultural frameworks centered on individualism and collectivism have been cited to explain such rhetorical divergence (Hofstede, 1984; Greenlee & Stück, 2004; Li & Andersen, 2017). Nevertheless, the present analysis demonstrates that purely cultural interpretations oversimplify cross-linguistic MDM variation, consistent with critical cautions raised in recent EAP and intercultural rhetoric research. Exclusive attribution to structural linguistic differences between Chinese and English is similarly incomplete, as this study does not directly measure cultural values or linguistic intentionality. Cultural and linguistic accounts therefore remain tentative contextual interpretations rather than definitive causal explanations. Future mixed-methods research that incorporates author interviews and ethnographic observation can extend the present findings by disclosing writers' conscious motivations behind MDM selection across linguistic and disciplinary contexts.

Cross-linguistic variation in overall MDM density and self-mention prevalence can be explained by genre conventions shaping Discussion section rhetoric. Academic genres evolve from recurrent communicative practices and reflect the shared expectations of disciplinary discourse communities (Swales, 1990; Bhatia, 1993). In English-medium research articles, the Discussion section prioritizes result interpretation, implication evaluation, and scholarly positioning within ongoing disciplinary debate (Basturkmen, 2012; Yang & Allison, 2003). These normative expectations encourage authors to emphasize the significance of findings and articulate explicit interpretive claims, directly accounting for the higher prevalence of interactional markers such as self-mentions in the English corpus. This functional alignment between Discussion rhetorical purposes and interactional MDM deployment corroborates Hyland's (2004, 2005b) taxonomy of MDMs, which positions self-mention as a core resource for constructing authorial identity and reinforcing argumentative credibility in interpretive academic writing. By contrast, Chinese-language articles published in domestic journals operate under local community norms that diverge substantially from international Q1 conventions, directly shaping the observed cross-linguistic and cross-contextual MDM patterns.

Genre-based academic training represents a further key factor underlying MDM variation. EAP genre pedagogy conventionally encourages explicit authorial stance formulation and persuasive interpretive writing in Discussion sections

(Swales, 1990). Researchers within this tradition naturally deploy self-mentions and stance markers more visibly, consistent with the higher MDM density observed in the English dataset. Writers trained primarily within domestic Chinese academic conventions favor impersonal constructions when presenting interpretations. This training-based divergence in rhetorical socialization extends comparative work on academic enculturation, demonstrating that disciplinary training systematically shapes not only general writing style but also the move-specific distribution of stance and engagement markers. Such differences are especially apparent in evaluative and interpretive rhetorical moves.

Institutional publishing conventions further reinforce these rhetorical regularities. International English journals require authors to report empirical outcomes, engage dialogically with readership, and explicitly articulate study contributions. These expectations prioritize stance and engagement markers, particularly in moves dedicated to result interpretation and significance evaluation (Hyland, 2001; Hyland & Tse, 2004; Mur-Dueñas, 2011). Domestic Chinese academic contexts favour impersonal, finding-centered presentation and downplay overt authorial visibility in interpretive commentary. This norm explains the lower frequency of authorial markers in Chinese texts and the greater reliance on impersonal structures and external citations to establish interpretive credibility. This finding closely aligns with prior observations of Chinese academic discourse, which consistently privileges implicit authorial presence and legitimizes claims through reference to prior scholarship rather than explicit self-argumentation (Li & Flowerdew, 2009; Huang, 2010; Liu, 2020). The present study confirms these established tendencies and extends them by showing that such preferences operate systematically across specific rhetorical moves rather than merely at the general textual level.

Taken together, these findings suggest that cross-linguistic differences in MDM use are likely to arise from the interaction of genre conventions, academic training, publication practices and linguistic differences, rather than from cultural values or linguistic factors alone. Recognizing this complexity helps avoid overly deterministic interpretations of intercultural differences in academic writing. Pedagogically, the results highlight the importance of contrastive genre instruction that draws attention to cross-linguistic differences in academic writing practices. Examining rhetorical patterns across languages can help novice writers understand how MDMs (such as self-mention and engagement) vary across academic contexts and publication environments. Chinese novice writers who aim for international publication may particularly benefit from explicit guidance on how and when to employ markers such as self-mentions, attitude markers and engagement markers in English Discussions. Furthermore, embedding corpus-based genre analysis into EAP or ESP curricula can provide concrete evidence of how MDMs

function across linguistic contexts, which in turn may further enhance writers' rhetorical flexibility and help them adapt their stance and engagement strategies to the expectations of international academic discourse.

Argumentative Progression Pattern

The pattern observed in the present study suggests that the deployment of MDMs is closely aligned with the evolving rhetorical moves in the Discussion section. As the argument develops across rhetorical moves, different types of markers become prominent in response to the communicative purposes of each move. In the initial moves of background information and reporting results, the frequent use of interactive markers such as transitions and frame markers reflects the need to organize the discourse and situate the findings within the broader research context. The present findings further confirm the claim made by Hyland (2005a) and Hyland & Tse (2004) that interactive organizational markers are fundamental to textual structuring and reader guidance in academic argumentation. When summarizing results, frame markers play an important role in signaling discourse boundaries and clarifying the organization of the discussion. This move-specific functional distribution further extends prior MDM research that has largely examined frame markers as general textual devices rather than resources strategically deployed to demarcate rhetorical stages within the Discussion section.

As the discussion progresses to commenting on results, the increased use of hedges and evidential markers suggests that writers tend to present explanations cautiously while linking their interpretations to previous studies. This observed preference directly corroborates the epistemic function of hedging outlined by Hyland (1998) and Salager-Meyer (1994), who emphasize that hedges enable academic writers to advance claims in tentative, negotiable terms to maintain scholarly objectivity. In the claim-making stage, the prominence of boosters and attitude markers indicates the salience of authorial stance as writers advance their arguments. The concentration of these stance markers within claim-oriented moves validates the functional framework proposed by Hyland (2005a) and Thompson (2001), and further refines their generalized model by demonstrating that boosters and attitude markers are not randomly distributed but concentrated specifically during interpretive claim formulation. Such stance resources enable writers to emphasize the significance of their findings and assert the validity of their interpretations.

Subsequently, the evaluative moves involve greater use of frame markers, self-mentions and boosters to foreground the contribution and significance of the study. This distribution aligns with early move structure analyses of Discussion sections by Hopkins & Dudley-Evans (1988) and Peacock

(2002), who identified evaluation and scholarly positioning as core rhetorical obligations in later Discussion moves. The use of self-mention particularly has been associated with the construction of authorial identity and the explicit claiming of research contributions (Hyland, 2004). The present results extend Hyland's (2004) account by locating identity construction and contribution claiming specifically within evaluative rhetorical moves, rather than treating self-mention as a diffuse textual feature. Finally, the frequent use of hedges and engagement markers in the move of drawing implications suggests that writers seek to present implications and involve readers in broader interpretations while maintaining a balanced stance by softening their tone. Engagement markers help establish a dialogic relationship with readers by directing their attention to the relevance and implications of the findings. Meanwhile, hedges are strategic devices to help writers soften their voice and build rapport with the audience (Hyland, 2005a). This coordinated use of hedges and engagement markers in implication-oriented moves confirms Hyland's (2005a) dialogic MDM framework and adds empirical specificity by mapping these interactional resources into the final argumentative stage of the Discussion section.

These findings illustrate that the sequential deployment of MDMs forms a coherent, data-driven interpretive pattern that functions as a key rhetorical resource facilitating the staged development of argumentation in the Discussion section. This move-linked progression of MDM use strongly confirms the linear rhetorical sequence outlined in prior move-analytic studies (Hopkins & Dudley-Evans, 1988; Peacock, 2002). Beyond confirming existing structural frameworks, the present study extends move analysis research by integrating MDM distribution into the sequential framework, which reveals how linguistic resources dynamically underpin each rhetorical transition across the Discussion.

The findings of this study also carry important pedagogical implications for the teaching of academic writing. The observed tendency in MDM deployment suggests that the use of MDMs is closely associated with the rhetorical purposes of different moves in the Discussion section. Therefore, academic writing instruction should not treat MDMs merely as isolated linguistic items, but rather as rhetorical resources that support the development of argumentation. This pedagogical recommendation builds on genre-based EAP studies and extends conventional MDM instruction, which often decontextualizes marker teaching from rhetorical purpose and discourse structure. In particular, teachers may guide students to recognize how different markers function across moves of the Discussion section. For example, interactive markers such as transitions and frame markers can help structure the opening of the discussion, hedges and evidentials can be used to present interpretations cautiously, boosters and attitude markers can strengthen claims, while engagement markers can help writers highlight implications

and involve readers in the concluding stage. This integrated move-MDM instructional approach offers a practical refinement to existing academic writing pedagogy and provides a structured template for novice writers to align MDM selection with rhetorical goals in each Discussion move. Therefore, integrating MDM instruction with move-based writing pedagogy may help novice writers develop more coherent and rhetorically effective Discussion sections.

CONCLUSION

This study investigated cross-disciplinary and cross-linguistic differences in the use of MDMs across rhetorical moves in the Discussion sections of Chinese and English research articles from soft and hard science fields. The findings suggest that MDM deployment is jointly shaped by disciplinary characteristics, linguistic contexts, genre conventions, academic training, and publication practices, which together shape their distribution across distinct rhetorical moves. The results extend prior research in genre analysis and intercultural rhetoric by demonstrating how MDMs are strategically employed within specific rhetorical moves of the Discussion section.

From a pedagogical perspective, the findings support genre-based EAP and ESP instructional approaches that integrate rhetorical move training and MDM instruction. Practical applications include: (1) linking specific MDM categories to corresponding rhetorical moves; (2) using annotated corpus examples to illustrate disciplinary and linguistic variations; and (3) training academic writers to make strategic, audience-aware rhetorical choices. Instruction targeting international publication can also help EFL writers align their rhetorical practices with global academic expectations, while remaining mindful of local discourse norms relevant to domestic publishing contexts.

Several limitations of the present study should be acknowledged. First, this research focuses only on four disciplines, which may restrict the generalizability of the findings to other academic fields. Second, the analysis is confined to published research articles within a limited time frame; accordingly, it does not account for MDM use in other academic genres nor capture the diachronic evolution of academic writing conventions. Third, the English corpus includes contributions by authors with diverse linguistic and institutional backgrounds. This heterogeneity may introduce confounding variables that exert unintended effects on the results. Accordingly, future research could broaden the disciplinary scope, examine alternative academic genres (e.g., conference presentations and grant proposals), conduct longitudinal investigations of novice writers' developmental trajectories, and analyze corpora with more strictly controlled author demographic profiles.

DECLARATION OF COMPETING INTEREST DISCLAIMER

None declared.

We disclose the use of ChatGPT for language polishing, stylistic suggestions, and rephrasing of the Introduction and Discussion sections to improve clarity and readability. ChatGPT was also used to identify and correct errors in spelling, number consistency, subject-verb agreement, article usage, word choice, and sentence coherence throughout the manuscript. No AI tools were used for data analysis, result generation, or interpretation. All conceptual decisions, analyses, and conclusions were made solely by the authors.

AUTHORS' CONTRIBUTIONS

Jingfen Jin: conceptualization; methodology; formal analysis; writing – original draft; writing – review & editing.

Angkana Tongpoon Patanasorn: supervision, conceptualization, writing – review & editing.

REFERENCES

- Ahmadi, L. (2022). Rhetorical structure of applied linguistics research article discussions: A comparative cross-cultural analysis. *Journal of Language and Education*, 8(3), 11-22. <https://orcid.org/0000-0003-0592-6059>
- Ahmed, A. A. (2023). The role of metadiscourse devices in Q1 Scopus-indexed linguistics research articles. *Journal of AlMaarif University College*, 34(1), 393-405. <https://doi.org/10.51345/v34i1.574.g348>
- Alharbi, L. M., & Swales, J. M. (2011). Arabic and English abstracts in bilingual language science journals: Same or different? *Languages in Contrast*, 11(1), 70-86. <https://doi.org/10.1075/lic.11.1.06alb>
- Amnuai, W. (2019). The textual organization of the discussion sections of accounting research articles. *Kasetsart Journal of Social Sciences*, 40(2), 389-394. <https://doi.org/10.1016/j.kjss.2017.10.007>
- Amnuai, W., & Wannaruk, A. (2012). Investigating move structure of English applied linguistics research article discussions published in International and Thai journals. *English Language Teaching*, 6(2), 1-13. <https://doi.org/10.5539/elt.v6n2p1>
- Ashofteh, Z., Shirvan, M. E., & Golparvar, E. (2020). The move structure of abstracts in applied linguistics research articles in light of the distribution and functions of metadiscourse markers. *Journal of Language and Linguistic Studies*, 16(4), 2077-2096. <https://doi.org/10.17263/jlls.851035>
- Azizah, N. (2017). *Code glosses in Discussion section of journal articles written by Indonesian and English native speakers* [Unpublished doctoral dissertation]. Universitas Negeri Jakarta. <http://repository.unj.ac.id/id/eprint/28751>
- Basturkmen, H. (2009). Commenting on results in published research articles and masters' dissertations in language teaching. *Journal of English for Academic Purposes*, 8(4), 241-251. <https://doi.org/10.1016/j.jeap.2009.07.001>
- Basturkmen, H. (2012). A genre-based investigation of discussion sections of research articles in Dentistry and disciplinary variation. *Journal of English for Academic Purposes*, 11(2), 134-144. <https://doi.org/10.1016/j.jeap.2011.10.004>
- Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines* (2nd ed.). Society for Research into Higher Education & Open University Press.
- Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. Longman.
- Biber, D., Ulla, C., & Thomas, U. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure*. John Benjamins Publishing.
- Biglan, A. (1973). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology*, 57(3), 195-203. <https://doi.org/10.1037/h0034701>
- Cao, F. (2014). *Metadiscourse in research articles: A comparative study across disciplines and research paradigms* [Unpublished doctoral dissertation]. National Institute of Education, Nanyang Technological University. <https://hdl.handle.net/10497/16725>
- Cao, F., & Hu, G. (2014). Interactive metadiscourse in research articles: A comparative study of paradigmatic and disciplinary influences. *Journal of Pragmatics*, 66, 15-31. <https://doi.org/10.1016/j.pragma.2014.02.007>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Crismore, A., Markkanen, R., & Steffensen, M. (1993). Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students. *Written Communication*, 10, 39-71. <https://doi.org/10.1177/0741088393010001002>
- De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Learning by generating vs. receiving instructional explanations: Two approaches to enhance attention cueing in animations. *Computers & Education*, 55(2), 681-691.

- Faqih, M. R. (2022). Boosters and self-mentions in Discussion section of English journal articles written by Indonesian-English native speakers [Unpublished doctoral dissertation]. Universitas Negeri Jakarta. <http://repository.unj.ac.id/id/eprint/24291>
- Gao, S., & Pramoolsook, I. (2022). Move-step structure of the Results and Discussion section of electronic Engineering research articles written by Chinese and Thai Researchers. *Journal of Teaching English for Specific and Academic Purposes*, 9(4), 725-739. <https://doi.org/10.22190/jtesap2104725g>
- Greenlee, D., & Stüeck, J. (2004). Individualist educators in a collectivist society: Insights from a cross-cultural model applied to China. *Missiology*, 32(4), 491-504. <https://doi.org/10.1177/009182960403200407>
- Hilmi, A. Z., & Afifi, N. (2021). A genre analysis on the discussion section of quantitative and qualitative research articles in ELT and linguistics. *JEELS (Journal of English Education and Linguistics Studies)*, 8(2), 341-369. <https://doi.org/10.30762/jeels.v8i2.3264>
- Hofstede, G. (1984). Culture's consequences: International differences in work-related values. *Sage*.
- Holmes, R. (1997). Genre analysis, and the social sciences: An Investigation of the structure of research article Discussion sections in three disciplines. *English for Specific Purposes*, 16(4), 321 -337. [https://doi.org/10.1016/S0889-4906\(96\)00038-5](https://doi.org/10.1016/S0889-4906(96)00038-5)
- Hopkins, A., & Dudley-Evans, T. (1988). A genre-based investigation of the discussion sections in articles and dissertations. *English for Specific Purposes*, 7(2), 113-121. [https://doi.org/10.1016/0889-4906\(88\)90029-4](https://doi.org/10.1016/0889-4906(88)90029-4)
- Huang, J. (2010). Chinese university students' perceptions of their difficulties in academic English writing. *Asian Social Science*, 6(11), 132-146. <https://doi.org/10.5539/ass.v6n11p132>
- Hu, G., & Cao, F. (2015). Disciplinary and paradigmatic influences on interactional metadiscourse in research articles. *English for Specific Purposes*, 39, 12-25. <https://doi.org/10.1016/j.esp.2015.03.002>
- Hyland, K. (1996a). Talking to the academy: Forms of hedging in science research articles. *Written Communication*, 13, 251-281. <https://doi.org/10.1177/0741088396013002004>
- Hyland, K. (1996b). Writing without conviction? Hedging in science research articles. *Applied Linguistics*, 17, 433-454. <https://doi.org/10.1093/applin/17.4.433>
- Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics*, 30(4), 437-455. [https://doi.org/10.1016/S0378-2166\(98\)00009-5](https://doi.org/10.1016/S0378-2166(98)00009-5)
- Hyland, K. (2001). Bringing in the reader: Addressee features in academic writing. *Written Communication*, 18(4), 549-574. <https://doi.org/10.1177/0741088301018004005>
- Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. University of Michigan Press.
- Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, 25(2), 156-177. <https://doi.org/10.1093/applin/25.2.156>
- Hyland, K. (2005a). *Metadiscourse: Exploring interaction in writing*. A&C Black.
- Hyland, K. (2005b). Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, 7(2), 173-192. <https://doi.org/10.1177/1461445605050365>
- Hyland, K. (2007). Applying a gloss: Exemplifying and reformulating in academic discourse. *Applied Linguistics*, 28(2), 266-285. <https://doi.org/10.1093/applin/28.2.266>
- Hyland, K. (2019). *Metadiscourse: Exploring interaction in writing*. Bloomsbury Publishing.
- Hyland, K., & Jiang, F. K. (2022). Metadiscourse choices in EAP: An intra-journal study of JEAP. *Journal of English for Academic Purposes*, 60, 101-166. <https://doi.org/10.1016/j.jeap.2022.101166>
- Indarti, D. (2022). Rhetorical moves and meta-discourse markers in Abstracts of Indonesian Scopus indexed journals across academic disciplines. *Journal of Languages and Language Teaching*, 10(4), 483-495. <https://doi.org/10.33394/jollt.v10i4.5173>
- Jalilifar, A., Hayati, A.M., & Namdari, N. (2012). A comparative study of research article Discussion sections of local and international applied linguistic journals. *The Journal of Asia TEFL*, 9(1), 1-29.
- Kadir, W. (2020). *The use of interactional metadiscourse markers in the Results and Discussion sections of Algerian scientific research articles* [Unpublished doctoral dissertation]. Universite Mouloud Mammaeri Tiziouzou.
- Kahkesh, M., & Alipour, M. (2017). Comparative analysis of metadiscourse markers in the Result and Discussion sections of Literature and Engineering research papers. *Iranian Journal of Applied Language Studies*, 9, 72-82. <https://doi.org/10.22111/ijals.2018.4192>
- Kustiyasari, D., Basthomi, Y., & Anugerahwati, M. (2021). Interactive and interactional metadiscourse markers in research articles of Indonesian expert writers. *Journal of English Educators Society*, 6(1), 90-95. <https://doi.org/10.21070/jees.v6i1.1082>

- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174. <https://doi.org/10.2307/2529310>
- Li, W., & Andersen, H. C. (2017). The complexity of indirect translation. *Orbis Litterarum*, 72(3), 181-208. <https://doi.org/10.1111/oli.12148>
- Liu, Y., & Buckingham, L. (2018). The schematic structure of discussion sections in applied linguistics and the distribution of metadiscourse markers. *Journal of English for Academic Purposes*, 34, 97-109. <https://doi.org/10.1016/j.jeap.2018.04.002>
- Li, Y., & Flowerdew, J. (2009). Academic discourse. In F. Sharifian (Ed.), *The Routledge handbook of language and culture* (pp. 465-478). Routledge.
- Loi, C.K., Evans, M. S., Lim, J.M., & Akkakoson, S. (2016). A comparison between Malay and English research article Discussions: A move analysis. *SAGE Open*, 6(2). <https://doi.org/10.1177/2158244016652925>
- Mohammad, A. S. N. (2012). A structural move analysis of Discussion sub-genre in applied linguistics. *Iranian Institute for Encyclopedia Research*, 2, 199-212. <https://works.bepress.com/nodushan/124/>
- Mu, C., Zhang, L. J., Ehrich, J., & Hong, H. (2015). The use of metadiscourse for knowledge construction in Chinese and English research articles. *Journal of English for Academic Purposes*, 20, 135-148. <https://doi.org/10.1016/j.jeap.2015.09.003>
- Mur-Dueñas, P. (2011). An intercultural analysis of metadiscourse features in research articles written in English and in Spanish. *Journal of Pragmatics*, 43(12), 3068-3079. <https://doi.org/10.1016/j.pragma.2011.05.002>
- Nguyen, T. T. L., & Pramoolsook, I. (2015). Move analysis of Results-Discussion chapters in TESOL master's theses written by Vietnamese students. *3L: The Southeast Asian Journal of English Language Studies*, 21(2), 1-15. <https://doi.org/10.17576/3l-2015-2102-01>
- Nugrahani, V. E. & Bram, B. (2020). Meta-discourse markers in scientific journal articles. *Langkawi: Journal of The Association for Arabic and English*, 6(1), 1-16. <https://doi.org/10.31332/lkw.v6i1.1528>
- Peacock, M. (2002). Communicative moves in the discussion section of research articles. *System*, 30(4), 479-497. [https://doi.org/10.1016/S0346-251X\(02\)00050-7](https://doi.org/10.1016/S0346-251X(02)00050-7)
- Safnil, A. (2013). A genre-based analysis on Discussion section of research articles in Indonesian written by Indonesian speakers. *International Journal of Linguistics*, 5(4), 50-70. <https://doi.org/10.5296/ijl.v5i4.3773>
- Salager-Meyer, F. (1994). Metadiscourse in academic writing: A comparative study of English and French research articles. *English for Specific Purposes*, 13(2), 149-170. [https://doi.org/10.1016/0889-4906\(94\)90003-8](https://doi.org/10.1016/0889-4906(94)90003-8)
- Soodmand A, H., Doosti, M., & Movassagh, H. (2018). A comparative study of generic structure of applied linguistics and chemistry research articles: The case of discussions. *Journal of Research in Applied Linguistics*, 9(1), 28-56. <https://doi.org/10.22055/raals.2018.13404>
- Sultan, M. A. (2011). Metadiscourse in Arabic and English research article abstracts. *Journal of Language and Social Psychology*, 30(4), 457-472. <https://doi.org/10.1177/0261927X11415558>
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Thompson, G. (2001). Interaction in academic writing: Learning to argue with the reader. *Applied Linguistics*, 22(1), 58-78. <https://doi.org/10.1093/applin/22.1.58>
- Tran, T. Q., & Duong, T. M. (2013). Hedging: A comparative study of research article results and discussion section in applied linguistics and chemical engineering. *English for Specific Purposes World*, 41(14), 1-13.
- Ulya, S. (2022). Rhetorical moves variations of research article discussion sections published in reputable journals. *Premise: Journal of English Education and Applied Linguistics*, 11(3), 632- 645. <https://doi.org/10.24127/pj.v11i3.5615>
- Vande Kopple, W.J. (1985). Some explanatory discourse on metadiscourse. *College Composition and Communication*, 36, 82-93.
- Vande Kopple, W. J. (2002). Metadiscourse, discourse, and issues in composition and rhetoric. In F. Barton & C. Stygall (Eds.), *Discourse studies in composition* (pp. 91-113). Hampton Press.
- Yang, R. Y., & Allison, D. (2003). Research articles in applied linguistics: moving from results to conclusions. *English for Specific Purposes*, 22(4), 365-385. [https://doi.org/10.1016/S0889-4906\(02\)00026-1](https://doi.org/10.1016/S0889-4906(02)00026-1)
- Yassaman, H. M., & Elham, Y. (2019). A genre-analysis of the Discussion section of Iranian and English ELT theses: A comparative study. *Theory and Practice in Language Studies*, 9(12), 1611-1623. <http://dx.doi.org/10.17507/tppls.0912.19>
- Zahra, S. D., Helen, T., & Ain, N. A. (2017). Rhetorical structure of integrated Results and Discussion chapter in master's dissertation across disciplines. *Discourse and Interaction*, 61-83. <https://doi.org/10.5817/DI2017-2-61>

APPENDIX A

English and Chinese Journals Selected for the Corpus

Discipline	English Journal	Chinese Journal
Hard Pure (HP) Physics	Fire Safety Journal	光学学报 (Acta Optica Sinica)
	Physica Medica	中国激光 (Chinese Journal of Lasers)
	Physics of the Earth and Planetary Interiors	原子核物理评论 (Nuclear Physics Review)
	PTEP (Progress of Theoretical and Experimental Physics)	波谱学杂志 (Chinese Journal of Magnetic Resonance)
	Computational Materials Science	中国科学:物理学 力学 天文学 Scientia Sinica (Physica, Mechanica & Astronomica)
Soft-Pure (SP) Linguistics	American Journal of Speech-Language Pathology	外语教学 (Foreign Language Teaching)
	English for Specific Purpose	现代外语 (Modern Foreign Language)
	Journal of Pragmatics	解放军外国语学院学报 (Journal of PLA University of Foreign Languages)
	Language and Communication Disorder	外语与外语教学 (Foreign languages and their teaching)
	Journal of Writing Research	中国外语 (Foreign Languages in China)
Hard-Applied (HA) Engineering	Applied Ergonomics	材料工程 (Journal of Materials Engineering)
	IATSS Research	材料研究学报 (Chinese Journal of Materials Research)
	Int. J. Human-Computer Sciences	采矿与安全工程学报 (Journal of Mining & Safety Engineering)
	Journal of Engineering Education	水利学报 (Journal of Hydraulic Engineering)
	Journal of Product Innovation Management	水力发电学报 (Journal of Hydroelectric Engineering)
Soft-Applied (SA) Education	Computer and Education	外语教学与研究 (Foreign Language Teaching and Research)
	Learning and Instruction	教育科学研究 (Educational Science Research)
	Language Learning	数学教育学报 (Journal of Mathematics Education)
	Studies in Higher Education	电化教育研究 (e-Education Research)
	System	远程教育杂志 (Journal of Distance Education)

APPENDIX B

A Combined Framework for Move Analysis in Discussion Section

Move	step	Code	Description
Move 1--- Background information		M1	The author restates the research purpose, research questions, research methodology or theoretical framework of the target research.
Move 2--- Reporting results	Step 1--- Expected results	M2S1	The author illustrates the expected results of the study, which directly answer the research questions referred to in previous texts.
	Step 2--- Unexpected results	M2S2	The author demonstrates the unexpected results, which do (not) target at or relate to the research questions directly, but are still produced by the research.
Move 3--- Summarizing results		M3	The author summarizes the results briefly, usually by integrating the sole results into one sentence.
Move 4--- Commenting on the results	Step 1--- Interpreting results	M4S1	The author interprets one of the results.
	Step 2--- Comparing results with previous studies	M4S2	The author compares the result with those reported in previous literature.
	Step 3--- Explaining results	M4S3	The author provides explanations of the (comparative) results, either similar or different. As an additional alternative, the author may turn to previous literature to support his/her explanation.
	Step 4--- Evaluating results	M4S4	The author makes an evaluation or indicates the significance of the results.
Move 5--- Making Claims		M5	The author offers an argument arising from the results generated in the target research (by referring to previous literature to support his/her argument).
Move 6--- Summarizing the study		M6	The author summarizes the research by combining/interpreting results as a whole.
Move 7--- Evaluating the study	Step 1--- Indicating the limitation of the study	M7S1	The author puts forward the possible limitations of the research so that future studies can get them refined.
	Step 2--- Indicating the significance of the study	M7S2	The author elucidates the importance/benefits of the study in relative field, disciplinary knowledge, practitioners or learners, etc.
	Step 3--- Evaluating generalizability/reliability	M7S3	The author makes comments on the generalizability or the reliability of the research.
Move 8--- <i>Deduction</i> from the study	Step 1--- Making suggestions for further study	M8S1	Suggestions are indicated for further research in order to have the current research improved.
	Step 2--- Drawing implications	M8S2	Pedagogical implications are clarified in terms of material development, instructors or learners.

APPENDIX C

An Analytical Framework of Metadiscourse Markers

Type	Subtype	Codes	Function	Example
<i>Interactive metadiscourse</i>				
Transitional marker	Additive	1ADD	to express relations of addition	in addition, furthermore, moreover
	Comparative	1COM	to express relations of comparison or contrast	similarly, however, in contrast
Frame marker	Referential	1REF	to express relations of cause and effect	thus, therefore, as a result
	Sequencers	2SEQ	to order discourse-internal units	first, second, finally
	Topicalizers	2TOP	to shift between topics	with regard to, concerning, turning to
Endophoric markers	Discourse-labels	2DIS	to label discourse stages	thus far, in sum, in brief
	Announcers	2ANN	to announce discourse goals	aim to, will, seek to
	Linear references	3LIN	to refer to the unfolding text	the next section, as noted earlier, in this paper
Evidential markers	Non-linear references	3NON	to refer to visual representations of the text	see Table 1, in Figure 2, as demonstrated in Excerpt 3
	Integral citations	4INT	to integrate the cited source into the text	according to X, as Y argued, in Z's study
Code glosses	Non-integral citations	4NON	to exclude the cited source from the text	"..." (X, 2013), ...previous research1, 2, 3
	Exemplifiers	5EXE	to elaborate meaning with examples	for example, for instance, e. g
	Reformulators	5REF	to rephrase a previous discourse unit	in other words, that is, i.e.
<i>Interactional metadiscourse</i>				
Hedges		6HED	to mitigate the degree of certainty or commitment	may, might, possible, perhaps, suggest, indicate
Boosters		7BOO	to increase the degree of certainty or commitment	will, demonstrate, show
Attitude markers		8ATT	to express affective attitudes or emotions	should, need to, interesting, surprisingly
Self-mentions		9SEL	to mark writer's explicit presence in text	I, we (exclusive), me, us, the author (researcher)my, our
Engagement markers	Directives	10DIR	to refer readers to actions	see, note, should, (cf.) (added by the author)
	Reader references	10REA	to make reference to readers	you, we (inclusive), the reader(s)
	Questions	10QUE	to anticipate readers' questions	rhetorical and real questions
	Knowledge appeals	10KOW	to make reference to shared knowledge)	well-known, obviously
	Personal asides	10PER	to address readers through interjections	---

APPENDIX D

Distribution of MDMs across groups (original counts, expected frequency and adjusted standardized residuals)

Appendix D1

Distribution of MDMs in Chinese soft vs. hard science (original counts, expected frequency and adjusted standardized residuals)

MDMs	M1				M2				M3				M4			
	Soft		Hard		Soft		Hard		Soft		Hard		Soft		Hard	
	O (E)	O (E)	ASR	ASR	O (E)	O (E)	ASR	ASR	O (E)	O (E)	ASR	ASR	O (E)	O (E)	ASR	ASR
Transition	3 (3.6)	14(13.4)	-0.4	0.4	52(48.1)	43(46.9)	0.8	-0.8	2(2.2)	0.5(0.8)	-0.2	0.2	201(183.4)	72(89.6)	2.4	-2.4
Frame marker	3 (4.5)	18(16.5)	-0.8	0.8	101(66.8)	31(65.2)	6.4*	-6.4*	4(5.0)	3(2.0)	-1	1	120(102.8)	33(50.2)	3	-3
Endophoric	3 (3.8)	15(14.2)	-0.5	0.5	27(83.0)	137(81.0)	-9.6*	9.6*	7(5.8)	0.5(2.2)	1.1	-1.1	41	112(50.2)	-10.9*	10.9*
Evidential	18 (15.6)	55(57.4)	0.8	-0.8	27(29.4)	31(28.6)	-0.6	0.6	2(2.2)	0.5(0.8)	-0.2	0.2	499(457.5)	182(223.5)	3.9	-3.9
Code gloss	3 (2.1)	7(7.9)	0.7	-0.70	15(13.7)	12(13.3)	0.5	-0.5	3(2.9)	1(1.1)	0.1	-0.1	104(93.4)	35(45.6)	2	-2
Hedge	4 (6.0)	24(22.0)	-1	1	67(83.6)	98(81.4)	-2.8	2.8	4(3.6)	0.5(1.4)	0.7	-0.7	449(443.4)	211(216.6)	0.5	-0.5
Booster	5 (3.2)	10(11.8)	1.2	-1.2	169(112.9)	54(110.1)	8.5*	-8.5*	5(4.3)	0.5(1.7)	0.7	-0.7	223	97(105.0)	1	-1
Attitude	0.5 (0.9)	3(3.1)	0.2	-0.2	14(8.6)	3(8.4)	2.6	-2.6	0.5(1.4)	0.5(0.6)	-0.7	0.7	31(29.6)	13(14.4)	0.5	-0.5
Self-mention	1 (2.1)	9 (7.9)	-0.9	0.9	10(6.1)	2(5.9)	2.3	-2.3	1(1.4)	0.5(0.6)	-0.7	0.7	43(32.2)	5(15.8)	3.3*	-3.3*
Engagement	12 (11.1)	40(40.9)	0.3	-0.3	35(64.8)	93(63.2)	-5.6*	5.6*	2(2.2)	0.5(0.8)	-0.2	0.2	178(229.1)	163(111.9)	-6.3*	6.3*

MDMs	M5				M6				M7				M8			
	Soft		Hard		Soft		Hard		Soft		Hard		Soft		Hard	
	O (E)	O (E)	ASR	ASR	O (E)	O (E)	ASR	ASR	O (E)	O (E)	ASR	ASR	O (E)	O (E)	ASR	ASR
Transition	31(32.9)	19(17.1)	-0.6	0.6	0.5(1.2)	2(1.8)	-0.3	0.3	10(11.1)	5(3.9)	-0.7	0.7	40(34.6)	20(25.4)	1.5	-1.5
Frame marker	6(5.9)	3(3.1)	0.1	-0.1	1(1.6)	3(2.4)	-0.7	0.7	10(11.1)	5(3.9)	-0.7	0.7	41(40.9)	30(30.1)	0	0
Endophoric	3(6.6)	7(3.4)	-2.4	2.4	1(8)	0.5(1.2)	0.3	-0.3	7(7.4)	3(2.6)	-0.3	0.3	4(7.5)	9(5.5)	-2	2
Evidential	24(18.4)	4(9.6)	2.4	-2.4	0.5(1.2)	2(1.8)	-0.3	0.3	10(8.1)	1(2.9)	1.4	-1.4	33(25.9)	12(19.1)	2.3	-2.3
Code gloss	1(1.3)	0.5(7.7)	-0.5	0.5	0.5(8)	0.5(1.2)	0.3	-0.3	1(1.5)	1(5)	-0.8	0.8	13(12.7)	9(9.3)	0.1	-0.1
Hedge	41(38.8)	18(20.2)	0.7	-0.7	0.5(1.6)	3(2.4)	-0.7	0.7	14(15.5)	7(5.5)	-0.8	0.8	34(47.8)	49(35.2)	-3.4*	3.4*
Booster	28(26.3)	12(13.7)	0.6	-0.6	2(1.2)	1(1.8)	1	-1	10(8.1)	0.5(2.9)	1.4	-1.4	10(11.0)	9(8.0)	-0.5	0.5
Attitude	10(10.5)	6(5.5)	-0.3	0.3	0.5(8)	1(1.2)	0.3	-0.3	10(9.6)	3(3.4)	0.3	-0.3	61(51.3)	28(37.7)	2.3	-2.3
Self-mention	6(5.9)	3(3.1)	0.1	-0.1	0.5(8)	1(1.2)	0.3	-0.3	0.5(2.2)	2(0.8)	-1.6	1.6	0.5(2.9)	4(2.1)	-1.7	1.7
Engagement	17(20.4)	14(10.6)	-1.4	1.4	0.5(8)	1(1.2)	0.3	-0.3	9(7.4)	1(2.6)	1.2	-1.2	12(14.4)	13(10.6)	-1	1

Notes: O = Observed frequency; E = Expected frequency; ASR = Adjusted Standardized Residual.
 p < .0025 (Bonferroni-corrected for 20 cells).
 * Indicates significant difference.

Appendix D2

Distribution of MDMs in English soft vs. hard science (original counts, expected frequency and adjusted standardized residuals)

MDMs	M1		M2		M3		M4									
	Soft	Hard	Soft	Hard	Soft	Hard	Soft	Hard								
	O (E)	O (E)	ASR	ASR	O (E)	O (E)	ASR	ASR								
Transition	14(16.9)	29(26.1)	-0.9	0.9	131(119.8)	74(85.2)	1.7	-1.7	0.5(1.3)	0.5(0.7)	-0.4	0.4	197(188.4)	121(129.6)	1	-1
Frame marker	31(27.9)	40(43.1)	0.8	-0.8	102(88.3)	49(62.7)	2.4	-2.4	6(4.4)	0.5(2.6)	1.4	-1.4	66(54.5)	26(37.5)	2.5	-2.5
Endophoric	2(12.2)	29(18.8)	-3.8*	3.8*	29(63.1)	79(44.9)	-6.9*	6.9*	0.5(1.3)	0.5(0.7)	-0.4	0.4	25(35.6)	35(24.4)	-2.8	2.8
Evidential	41(42.4)	67(65.6)	-0.3	0.3	48(54.9)	46(39.1)	-1.5	1.5	1(1.3)	0.5(0.7)	-0.4	0.4	507(481.8)	306(331.2)	2.1	-2.1
Code gloss	7(7.1)	11(10.9)	0	0	20(18.7)	12(13.3)	0.5	-0.5	0.5(1.3)	0.5(0.7)	-0.4	0.4	69(62.2)	36(42.8)	1.4	-1.4
Hedge	42(42.0)	65(65.0)	0	0	133(141.5)	109(100.5)	-1.2	1.2	2(1.9)	0.5(1.1)	0.1	-0.1	516(539.2)	394(370.8)	-1.8	1.8
Booster	8(7.1)	10(10.9)	0.5	-0.5	230(225.6)	156(160.4)	0.5	-0.5	2(1.9)	0.5(1.1)	0.1	-0.1	162(164.1)	115(112.9)	-0.3	0.3
Attitude	5(5.5)	9(8.5)	-0.3	0.3	36(32.7)	20(23.3)	0.9	-0.9	0.5(1.3)	0.5(0.7)	-0.4	0.4	62(80.6)	74(55.4)	-3.3*	3.3*
Self-mention	71(64.4)	93(99.6)	1.2	-1.2	155(147.9)	98(105.1)	1	-1	1(1.3)	0.5(0.7)	-0.4	0.4	160(167.1)	122(114.9)	-0.9	0.9
Engagement	60(55.7)	82(86.3)	0.8	-0.8	177(168.4)	111(119.6)	1.1	-1.1	1(1.3)	0.5(0.7)	-0.4	0.4	201(191.4)	122(131.6)	1.1	-1.1

MDMs	M5		M6		M7		M8									
	Soft	Hard	Soft	Hard	Soft	Hard	Soft	Hard								
	O (E)	O (E)	ASR	ASR	O (E)	O (E)	ASR	ASR								
Transition	41(38.7)	48(50.3)	0.5	-0.5	1(0.8)	2(2.2)	0.3	-0.3	55(46.6)	70(78.4)	1.6	-1.6	50(51.4)	69(67.6)	-0.3	0.3
Frame marker	14(11.3)	12(14.7)	1.1	-1.1	3(2.3)	6(6.7)	2.3	-2.3	31(27.6)	43(46.4)	0.8	-0.8	25(21.2)	24(27.8)	1.1	-1.1
Endophoric	4(3.9)	5(5.1)	0.1	-0.1	0.5(1.5)	5(4.5)	-0.5	0.5	3(3.7)	7(6.3)	-0.5	0.5	2(7.3)	15(9.7)	-2.6	2.6
Evidential	24(30.9)	47(40.1)	-1.7	1.7	1(3.1)	11(8.9)	-1.5	1.5	49(63.7)	122(107.3)	-2.5	2.5	63(72.5)	105(95.5)	-1.6	1.6
Code gloss	13(10.0)	10(13.0)	1.3	-1.3	2(1.0)	2(3.0)	1.1	-1.1	21(19.0)	30(32.0)	0.6	-0.6	23(26.8)	39(35.2)	-1	1
Hedge	85(86.6)	114(112.4)	-0.3	0.3	2(2.1)	6(5.9)	0	0	77(82.8)	145(139.2)	-0.9	0.9	140(142.0)	189(187.0)	-0.3	0.3
Booster	21(27.4)	42(35.6)	-1.7	1.7	2(2.3)	7(6.7)	-0.3	0.3	36(38.4)	67(64.6)	-0.5	0.5	36(42.7)	63(56.3)	-1.4	1.4
Attitude	35(37.0)	50(48.0)	-0.5	0.5	3(1.3)	2(3.7)	1.8	-1.8	19(15.7)	23(26.3)	1.1	-1.1	98(73.8)	73(97.2)	4*	-4*
Self-mention	26(20.0)	20(26.0)	1.8	-1.8	1(2.6)	9(7.4)	-1.2	1.2	93(90.2)	149(151.8)	0.4	-0.4	52(56.5)	79(74.5)	-0.8	0.8
Engagement	29(26.1)	31(33.9)	0.8	-0.8	2(1.0)	2(3.0)	1.1	-1.1	35(31.3)	49(52.7)	0.9	-0.9	51(45.8)	55(60.2)	1.1	-1.1

Notes: O = Observed frequency; E = Expected frequency; ASR = Adjusted Standardized Residual.
 p < .0025 (Bonferroni-corrected for 20 cells).
 * Indicates significant difference.

Appendix D3

Distribution of MDMs in Chinese vs. English soft science (original counts, expected frequency and adjusted standardized residuals)

	M1				M2				M3				M4			
	CH		EN		CH		EN		CH		EN		CH		EN	
	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR
Transition	3(2.7)	0.2	14(14.3)	-0.2	52(60.0)	1.3	131(123.0)	-1.3	0.5(0.8)	0.5(1.2)	0.3	-0.3	201(195.1)	197(202.9)	0.6	-0.6
Frame marker	3(5.4)	-1.2	31(28.6)	1.2	101(66.5)	-5.5*	102(136.5)	5.5*	3(3.7)	6(5.3)	-0.6	0.6	120(91.2)	66(94.8)	4.3*	-4.3*
Endophoric	3(0.8)	2.7	2(4.2)	-2.7	27(18.3)	2.5	29(27.7)	-2.5	0.5(0.8)	0.5(1.2)	0.3	-0.3	41(32.3)	25(33.7)	2.1	-2.1
Evidential	18(9.4)	3.4*	41(49.6)	-3.4*	27(24.6)	0.6	48(50.4)	-0.6	0.5(0.8)	1(1.2)	0.3	-0.3	499(493.1)	507(512.9)	0.4	-0.4
Code gloss	3(1.6)	1.2	7(8.4)	-1.2	15(11.5)	1.3	20(23.5)	-1.3	1(0.8)	0.5(1.2)	0.3	-0.3	104(84.8)	69(88.2)	3	-3
Hedge	4(7.3)	-1.4	42(38.7)	1.4	67(65.5)	0.2	133(134.5)	-0.2	0.5(1.2)	2(1.8)	-0.3	0.3	449(473.0)	516(492.0)	-1.8	1.8
Booster	5(2.1)	2.3	8(10.9)	-2.3	169(130.7)	-4.7*	230(268.3)	4.7*	0.5(1.2)	2(1.8)	-0.3	0.3	223(188.7)	162(196.3)	3.7*	-3.7*
Attitude	0.5(1.0)	0.1	5(5.0)	-0.1	14(16.4)	0.7	36(33.6)	-0.7	0.5(0.8)	0.5(1.2)	0.3	-0.3	31(45.6)	62(47.4)	-3.1	3.1
Self-mention	1(11.4)	-3.8*	7(60.6)	3.8*	10(54.1)	-7.7*	155(110.9)	7.7*	0.5(0.8)	1(1.2)	0.3	-0.3	43(99.5)	160(103.5)	-8.1*	8.1*
Engagement	12(11.4)	0.2	60(60.6)	-0.2	35(69.5)	-5.4*	177(142.5)	5.4*	0.5(0.8)	1(1.2)	0.3	-0.3	178(185.8)	201(193.2)	-0.8	0.8

	M5				M6				M7				M8			
	CH		EN		CH		EN		CH		EN		CH		EN	
	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR
Transition	31(26.2)	1.3	41(45.8)	-1.3	0.5(0.8)	0.4	1(1.2)	-0.4	10(10.6)	55(54.4)	-0.2	0.2	40(28.4)	50(61.6)	2.8	-2.8
Frame marker	6(7.3)	-0.6	14(12.7)	0.6	1(1.5)	-0.6	3(2.5)	0.6	10(6.7)	31(34.3)	1.4	-1.4	41(20.8)	25(45.2)	5.6*	-5.6*
Endophoric	3(2.5)	0.4	4(4.5)	-0.4	1(0.8)	0.4	0.5(1.2)	-0.4	7(1.6)	3(8.4)	4.6*	-4.6*	4(1.9)	2(4.1)	1.9	-1.9
Evidential	24(17.5)	2.1	24(30.5)	-2.1	0.5(0.8)	0.4	1(1.2)	-0.4	10(9.7)	49(49.3)	0.1	-0.1	33(30.3)	63(65.7)	0.6	-0.6
Code gloss	1(5.1)	-2.3	13(8.9)	2.3	0.5(1.1)	-0.2	2(1.9)	0.2	1(3.6)	21(18.4)	-1.5	1.5	13(11.4)	23(24.6)	0.6	-0.6
Hedge	41(45.8)	-1.1	85(80.2)	1.1	0.5(1.1)	-0.2	2(1.9)	0.2	14(14.9)	77(76.1)	-0.3	0.3	34(54.9)	140(119.1)	-3.9*	3.9*
Booster	28(17.8)	3.2*	21(31.2)	-3.2*	2(1.5)	0.5	2(2.5)	-0.5	10(7.5)	36(38.5)	1	-1	10(14.5)	36(31.5)	-1.5	1.5
Attitude	10(16.4)	-2.1	35(28.6)	2.1	0.5(1.5)	-0.6	3(2.5)	0.6	10(4.7)	19(24.3)	2.7	-2.7	61(50.2)	98(108.8)	2.1	-2.1
Self-mention	6(11.6)	-2.1	26(20.4)	2.1	0.5(0.8)	0.4	1(1.2)	-0.4	0.5(15.1)	93(78.6)	-4.4*	4.4*	0.5(16.7)	52(36.3)	-4.8*	4.8*
Engagement	17(16.7)	0.1	29(29.3)	-0.1	0.5(1.1)	-0.2	2(1.9)	0.2	9(7.2)	35(36.8)	0.8	-0.8	12(19.9)	51(43.1)	-2.2	2.2

Notes: O = Observed frequency; E = Expected frequency; ASR = Adjusted Standardized Residual.
 p < .0025 (Bonferroni-corrected for 20 cells).
 * Indicates significant difference.

Appendix D4

Distribution of MDMs in Chinese vs. English hard science (original counts, expected frequency and adjusted standardized residuals)

	M1				M2				M3				M4			
	CH		EN		CH		EN		CH		EN		CH		EN	
	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR
Transition	14(13.3)	29(29.7)	0.2	-0.2	43(46.9)	74(70.1)	-0.8	0.8	2(2.3)	0.5(0.7)	-0.4	0.4	72(43.8)	36(64.2)	5.7*	-5.7*
Frame marker	18(18.0)	40(40.0)	0	0	31(32.1)	49(47.9)	-0.2	0.2	4(3.8)	0.5(1.2)	0.2	-0.2	33(62.5)	121(91.5)	-5*	5*
Endophoric	15(13.6)	29(30.4)	0.5	-0.5	137(86.5)	79(129.5)	7.7*	-7.7*	7(6.0)	0.5(2.0)	0.9	-0.9	112(56.0)	26(82.0)	10*	-10*
Evidential	55(37.8)	67(84.2)	3.8*	-3.8*	31(30.8)	46(46.2)	0	0	-2.3	0.5(0.7)	-0.4	0.4	182(88.1)	35(128.9)	13.7*	-13.7*
Code gloss	7(5.6)	11(12.4)	0.7	-0.7	12(9.6)	12(14.4)	1	-1	3(3.0)	0.5(1.0)	0	0	35(138.4)	306(202.6)	-12.4*	12.4*
Hedge	24(27.5)	65(61.5)	-0.9	0.9	98(82.9)	109(124.1)	2.3	-2.3	4(3.8)	0.5(1.2)	0.2	-0.2	211(245.6)	394(359.4)	-3.3*	3.3*
Booster	10(6.2)	10(13.8)	1.9	-1.9	54(84.1)	156(125.9)	-4.6*	4.6*	5(4.5)	0.5(1.5)	0.5	-0.5	97(86.0)	115(126.0)	1.6	-1.6
Attitude	3(3.7)	9(8.3)	-0.5	0.5	3(9.2)	20(13.8)	-2.7	2.7	0.5(1.5)	0.5(0.5)	-0.9	0.9	13(35.3)	74(51.7)	-5*	5*
Self-mention	9(31.6)	93(70.4)	-5.3*	5.3*	2(40.1)	98(59.9)	-8.1*	8.1*	1(1.5)	0.5(0.5)	-0.9	0.9	5(51.5)	122(75.5)	-8.7*	8.7*
Engagement	40(37.8)	82(84.2)	0.5	-0.5	93(81.7)	111(122.3)	1.8	-1.8	2(2.3)	0.5(0.7)	-0.4	0.4	163(115.7)	122(169.3)	6.1*	-6.1*

	M5				M6				M7				M8			
	CH		EN		CH		EN		CH		EN		CH		EN	
	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR	O (E)	ASR
Transition	19(12.5)	48(54.5)	2.2	-2.2	2(0.9)	2(3.1)	1.3	-1.3	5(3.0)	70(72.0)	1.3	-1.3	20(18.2)	69(70.8)	0.5	-0.5
Frame marker	3(2.8)	12(12.2)	0.2	-0.2	3(2.1)	6(6.9)	0.7	-0.7	5(1.9)	43(46.1)	2.4	-2.4	30(11.1)	24(42.9)	6.6*	-6.6*
Endophoric	7(2.2)	5(9.8)	3.6*	-3.6*	0.5(1.4)	5(4.6)	-0.4	0.4	3(0.4)	7(9.6)	4.3*	-4.3*	9(4.9)	15(19.1)	2.1	-2.1
Evidential	4(9.5)	47(41.5)	-2.1	2.1	2(3.1)	11(9.9)	-0.8	0.8	1(4.9)	122(118.1)	-2	2	12(23.9)	105(93.1)	-2.9	2.9
Code gloss	0.5(2.1)	10(8.9)	-0.8	0.8	0.5(0.7)	2(2.3)	0.4	-0.4	1(1.2)	30(29.8)	-0.2	0.2	9(9.8)	39(38.2)	-0.3	0.3
Hedge	18(24.6)	114(107.4)	-1.8	1.8	3(2.1)	6(6.9)	0.7	-0.7	7(6.0)	145(146.00)	0.5	-0.5	49(48.7)	189(189.3)	0.1	-0.1
Booster	12(10.1)	42(43.9)	0.7	-0.7	1(1.9)	7(6.1)	-0.8	0.8	0.5(2.7)	67(65.3)	-1.1	1.1	9(14.7)	63(57.3)	-1.7	1.7
Attitude	6(10.5)	50(45.5)	-1.6	1.6	1(0.7)	2(2.3)	0.4	-0.4	3(1.0)	23(25.0)	2	-2	28(20.7)	73(80.3)	1.9	-1.9
Self-mention	3(4.3)	20(18.7)	-0.7	0.7	1(2.4)	9(7.6)	-1.1	1.1	2(6.0)	149(145.0)	-1.9	1.9	4(17.0)	79(66.0)	3.7*	-3.7*
Engagement	14(8.4)	31(36.6)	2.3	-2.3	1(0.7)	2(2.3)	0.4	-0.4	1(2.0)	49(48.0)	-0.7	0.7	13(13.9)	55(54.1)	-0.3	0.3

Notes: O = Observed frequency; E = Expected frequency; ASR = Adjusted Standardized Residual.
 p < .0025 (Bonferroni-corrected for 20 cells).
 * Indicates significant difference.