Hedging use in Spanish academic writing: A contrastive corpus-based study

Gang Yao & Wei Sun
Beijing Foreign Studies University (China)
g.yao@bfsu.edu.cn, felipe@bfsu.edu.cn

Abstract

Hedging, a crucial rhetorical strategy in academic discourse, allows writers to express appropriate caution and modesty while presenting claims. While extensively studied in English academic writing, research on hedging in Spanish academic discourse, particularly comparing non-native and native writers, remains limited. This study investigates hedging practices in Spanish academic writing across three writer groups: non-native (L1 Chinese) novice, native novice, and native expert writers. Using a corpus-based approach, it examines the distribution of lexico-grammatical and functional hedging categories, individual hedging preferences, and specific hedging functions. The analysis reveals significant differences in the use of modal verb hedges and shield hedges, with non-native writers employing these devices less frequently than both native novice and expert writers. Non-native writers also demonstrate a distinct overuse of certain approximators and informal expressions, while underutilizing the conditional morpheme *-ría*. Functionally, non-native writers show a preference for less tentative strategies and overt personal involvement, contrasting with the more detached or nuanced stance-taking of native writers. These variations may be attributed to linguistic challenges, cultural influences, and L2 proficiency factors. The study highlights the complex interplay between cultural rhetorical traditions, language expertise, and pragmatic competence in academic hedging. Pedagogical implications include the need for explicit instruction in pragmatic functions of hedging devices, raising awareness of cross-cultural differences in authorial presence, and enhancing register awareness among L2 writers. This research contributes to our understanding of hedging in Spanish academic discourse and informs L2 writing instruction in academic contexts.

Keywords: Hedging, academic writing, Spanish for academic purposes, L2 writing.

RESUMEN

El uso de la atenuación en la escritura académica en español: Estudio contrastivo basado en corpus

La atenuación es una estrategia retórica crucial en el discurso académico que permite a los escritores expresar la debida cautela y modestia al presentar sus argumentos. Si bien se ha estudiado ampliamente en la escritura académica en inglés, sigue siendo limitada la investigación sobre la atenuación en el discurso académico en español, especialmente en lo que respecta a la comparación entre escritores no nativos y nativos. Este artículo investiga las prácticas de atenuación en la escritura académica en español en tres grupos de escritores: escritores noveles no nativos (L1 chino), noveles nativos y expertos nativos. Mediante un enfoque basado en corpus, se examina la distribución de las categorías lexicogramaticales y funcionales de la atenuación, las preferencias individuales de atenuación y sus funciones específicas. El análisis revela diferencias significativas en el uso de verbos modales y *shields*, ya que los escritores no nativos emplean estos recursos con menos frecuencia que los escritores nativos principiantes y expertos. Los escritores no nativos también muestran un uso excesivo de ciertos aproximadores y expresiones informales, mientras que infrautilizan el morfema condicional *-ría*. Funcionalmente, los escritores no nativos muestran una preferencia por estrategias menos tentativas y una implicación personal manifiesta, en contraste con la adopción de posturas más distantes o matizadas de los escritores nativos. Estas variaciones pueden atribuirse a dificultades lingüísticas, influencias culturales y nivel de dominio de la L2. El estudio destaca la compleja interacción entre las tradiciones retóricas culturales, la experiencia lingüística y la competencia pragmática en la atenuación académica. Entre las implicaciones pedagógicas se señala la necesidad de una instrucción explícita sobre las funciones pragmáticas de los mecanismos de atenuación, así como la sensibilización sobre las diferencias interculturales en la presencia del autor, así como sobre el uso del registro entre los escritores de L2. Esta investigación arroja luz sobre el comportamiento de la atenuación en el discurso académico español y sobre la enseñanza de la escritura en L2 en contextos académicos.

Palabras clave: escritura académica, español para fines específicos, escritura en L2.

1. Introduction

Academic writing, a form of written discourse intended to contribute to the creation and dissemination of scholarly knowledge, is arguably one of the most crucial aspects of professional development for students and lecturers.

Admittedly, the knowledge created and disseminated through academic writing primarily concerns propositional content (e.g., *Men have a shorter lifespan than women*). However, many researchers share the point that academic texts are not only about transmitting content or information objectively but also aim to convince and influence the audience (Hyland, 1998; Markkanen & Schröder, 1997; Martín-Martín, 2008; Vold, 2006). The latter aspect concerns the writer's pragmatic competence, the ability to deliver his/her intended message with all its nuances in any socio-cultural settings and to produce appropriate texts that meet rhetorical conventions and expectations of their discourse community (Chen & Zhang, 2017; Fraser, 2010) (e.g., *The results seem to suggest that men have a shorter lifespan than women*).

One of the most important indicators of the writer's pragmatic competence is the technique of hedging. The concept of *hedging* originates in logic and semantics (Lakoff, 1973; see Schröder & Zimmer, 1997) but has since been further developed in the areas of pragmatics and discourse analysis. In the meantime, the concept itself has become abundant to the point that it has overlapped with other related concepts such as epistemic modality and evidentiality (see Markkanen & Schröder, 1997; Martín-Martín, 2008) and varying views have formed from domains such as mitigation (Albelda & Estellés, 2021; Caffi, 1999; Fraser, 1980), politeness (Brown & Levinson, 1987; Myers, 1989), and vagueness (Channell, 1985, 1994). As a result of this development, however, the notion becomes unclear and even appears to "have reached a state of definitional chaos" (Markkanen & Schröder, 1997, p. 15). While there are many scholars who tried to define hedging from different perspectives, they generally agree that hedging refers to the writer's rhetorical strategy of using certain explicit linguistic means (so-called *hedges,* such as *may, seem, probably, it is likely that*) to withhold complete commitment to the truth of a proposition or to modulate the force of their expressions (Crompton, 1997; Fraser, 2010; Hinkel, 2005; Hyland, 1996b, 1998, 2005a; Mur-Dueñas, 2021).

Hedges have been pointed to as an omnipresent feature in academic discourse, in which their proper use is vital. This is because they allow writers to convey possibility and tentativeness, to take a stance on their statements, and to engage their readers in a dialogic space (Hyland, 1996b; Mur-Dueñas, 2021). First, it is necessary for writers to express tentativeness and possibility as the indeterminate nature of knowledge in academic genres requires that all claims conveying new knowledge should be hedged (Mur-Dueñas, 2021; Myers, 1989; Vold, 2006). Second, it is also common for writers to adopt a

position on statements because not all information is presented as an accredited fact. That is to say, sometimes a statement is displayed as an opinion, which writers tend to hedge in order to show a lack of full commitment (Hyland, 2005b). Third, it seems reasonable for writers to create an imaginary dialogue with their peer readers, where the latter might have alternative interpretations or opinions so writers use hedges to anticipate their potential criticism and to negotiate with them, thereby reducing the negatability and threats of claims (Meyer, 1997) and protecting their face (Albelda, 2018; Figueras, 2018; Martín-Martín, 2008). Therefore, the proper use of hedging can help writers not only achieve their communicative goals but also gain credibility and acceptance in the discourse community to which they belong; on the other hand, however, failing to hedge appropriately could lead to miscommunication and misunderstanding between the writer and the reader (Abdollahzadeh, 2019; Fraser, 2010; Yang, 2013).

Since hedging is a common and useful technique in academic discourse communication, it is obvious that it should be studied and understood thoroughly and should also be taught in classrooms. A complete examination of hedging use can contribute to our better understanding of how academic writers employ evidential reasoning and structure argumentation in a nuanced and balanced manner (Hyland, 1996a), thereby helping student writers adhere to the academic conventions of the field. The more we understand its features, the more we can assist the students (both native and non-native) in their studies and professional careers.

In what follows, we will briefly review previous studies on hedging use in academic writing.

2. Hedging use in academic writing: a review

2.1. Different perspectives on hedging

As discussed in the introduction, hedging is an essential pragmatic strategy in academic writing that allows writers to carefully qualify their claims and open a dialogue with readers. Through strategic use of hedging devices, scholars can put forth, for instance, novel assertions and counterarguments while still appearing objective and humble. Given the vital role of hedging devices in scholarly argumentation and persuasion, numerous perspectives have been adopted to further our understanding of this linguistic

phenomenon. A growing body of research has looked into hedging through a cross-linguistic/cultural, cross-disciplinary, cross-expertise, cross-generic and diachronic lens.

A predominant research focus has centered on cross-linguistic/cultural variations in academic hedging practices, with many contrastive studies (Abdollahzadeh, 2019; Akbas, 2012; Dontcheva-Navratilova, 2016; Hu & Cao, 2011; Mur-Dueñas, 2021; Vassileva, 1997, 2001; Vold, 2006; Yang, 2013) pointing to more extensive and frequent use of hedges in English academic writing compared to some other languages. In the second language setting, English L2 writers were found to use fewer and restricted hedging devices than the native speakers (Hinkel, 2005). However, a few studies have reported inconsistent findings. Clyne (1994), for instance, found that German writers used far more hedges in both German and English texts than their English-speaking peers. Similarly, Carrió-Pastor's (2016) study indicates that Spanish-speaking writers used more hedges in English papers than the native-speaking cohorts. The study by Alonso Alonso (2019) reported that both Spanish L1 and English L2 writers employed more hedging resources than English L1 writers. According to Clyne (1994) and Hu and Cao (2011), factors like cultural preferences for detachment vs involvement, rhetorical norms, epistemic beliefs about science, and reader vs writer responsibility likely contribute to the variations of hedging use.

Hedging expressions are naturally assumed to display disciplinary variations, as it is reasonable that different fields may have differing conventions of certainty and bases of argumentation (Markkanen & Schröder, 1997). However, the differences might not be as clear as has been supposed. On the one hand, in Abdi (2002), Hyland (2005b), and Carrió-Pastor (2019) hedges were found considerably more frequent in soft knowledge disciplines like philosophy and linguistics compared to hard knowledge sciences like physics and engineering. On the other hand, however, Vold (2006) and Sanjaya (2015) only detected negligible cross-field differences in hedging use.

A few studies investigate the variations of hedging use across academic genres. For instance, Salager-Meyer (1997) documented that medical editorials and review articles were more heavily hedged in comparison with research papers and case reports, as these genres are different in the extent to which the authors intend to make pretensions to generalization in their writing. Varttala (2001) found a different use of hedging in research articles vs. popular scientific articles, due to the different audiences for which these

two genres are intended. More studies (Hyland, 1998; Martín-Martín, 2008; Salager-Meyer, 1994, 1997) are interested in hedging use in so-called "part-genres" (Dudley-Evans, 2000, p. 5), which examine the hedges across different rhetorical sections of a scientific paper. They found that hedging was particularly common in research article introductions and discussions where claims tend to be more tentative and speculative. Method and results sections, on the contrary, commonly use unhedged language as confirmatory statements are the rule in these less discursive and commentative sections.

Some studies (Aull & Lancaster, 2014; Smirnova & Strinyuk, 2020) have found that novice authors hedge their claims to a lesser degree than expert writers, suggesting hedging can vary across expertise and experience. Hedges are used strategically by adept authors to navigate controversy, express cautiousness, and maintain relations with readers; in contrast, novice writers, as evidenced in those studies, often encounter some difficulties in achieving such communicative goals.

Another perspective that sparks research interest is the diachronic change of hedging in academic writing. For example, Poole et al. (2019) showed an overall decreasing trend of using hedges over time in the biochemical research field. In a similar vein, Yao et al. (2023) found that the use of hedges had significantly decreased in the past 25 years in the research articles of the journal *Science*, a trend that, they believe, was in line with the rising "linguistic positivity" phenomenon in academic writing. However, in the diachronic study of stance by Hyland and Jiang (2016), there were dramatic falls in hedging use in soft knowledge disciplines but steady increases in hard sciences over the past 50 years, suggesting that diachronic analyses of hedging practices may require nuanced interpretation with knowledge communities.

A small group of researchers focused on the less-explored aspects of hedging use. For example, Liu and Tseng (2021) were interested in the paradigmatic influence on the varied use of hedging in texts, and followed two qualitative approaches, i.e., narrative inquiry and grounded theory. Vold (2006) also looked at the possible gender effect on hedging in academic prose, although her study did not prove the existence of such an association.

2.2. Hedging use in Spanish academic writing

As can be seen, current scholarship primarily concentrates on hedging patterns displayed in English contexts, or with English being the main language contrasted. Other languages seem to have received scant attention,

which is clearly associated with the 'Anglocentrism' in the international academic community. In the Spanish-speaking academic sphere, there exists a robust line of research on hedging, though it has predominantly focused on oral discourse and colloquial conversation (Albelda et al., 2014; Albelda, 2020; Briz, 1995; Cestero & Albelda, 2020, to name a few). While these oral discourse studies have contributed significantly to our understanding of hedging mechanisms in Spanish, our focus remains on the written academic domain, where some local studies have examined hedging use in Spanish academic writing.

A targeted review reveals that these studies seem to have particular interest in the genres of student writing. For example, Acín Villa's (2016) working paper provides an initial approach to the study of hedging in the conclusion chapter of PhD dissertations in Spanish and examines how authors use different techniques to soften their claims. The main findings are that they use a limited number of procedures to attenuate their claims and that rhetorical and academic conventions play an important role in shaping the way information is presented in doctoral theses.

Also based on student writing genres, a few studies have shifted their attention to metadiscursive variations in general across discipline, part-genre, and gender. Navarro et al. (2022), for instance, analyze the use of metadiscourse in undergraduate dissertations across engineering and humanities. Their study found that undergraduate students in both disciplines tend to mitigate and boost claims in their dissertations, but humanities students use almost double the relative number of hedges and boosters used by engineering cohorts in introductions and conclusions. On the other hand, Cañada Pujols and Bach (2022) explore the use of metadiscursive markers in undergraduate thesis abstracts written by Spanish university students, finding that hedges are present in all rhetorical moves of the abstracts, but their frequency is higher in the presentation of objectives and results. Finally, Núñez-Román et al. (2021) analyze the use of stance markers in 187 undergraduate dissertations in education. They detect a gender effect on students' academic writing. For example, in terms of hedging use, female student writers employed more hedges than their male peers, suggesting they have more awareness of adopting a tentative stance in the discourse community.

Focusing on more specific hedging structures, Chao Parapar (2018) explores the use of verbs *poder*, *caber*, *deber*, *haber*, and *convenir* followed by the infinitive

in Spanish scientific articles from four disciplines. This study found that the most productive hedging construction is *poder* + infinitive and that technical disciplines such as engineering use these hedging resources less than other disciplines.

It is clear from these available studies that hedging is also a prevalent feature in Spanish academic discourse, with variations arising from factors such as the discipline of study, the rhetorical section of the text, and the writer's gender. However, a review of these studies also reveals a gap in our understanding of how hedging use varies (or not) across different lingua-cultural backgrounds; specifically, how it is used by native versus non-native Spanish writers. This is a critical aspect of Spanish academic writing that warrants further investigation given the potential influence of linguistic and cultural background on hedging behaviors, as demonstrated in many English studies.

The purpose of the current study, therefore, is to explore how Spanish L2 (non-native; NNS) and L1 (native; NS) writers employ hedging devices in their academic writing. Specifically, it seeks to provide answers to the following research questions through a contrastive corpus analysis:

RQ1: What hedging categories are employed by NNS vs. NS Spanish academic writers? Are there any significant differences?

RQ2: What individual hedging devices are preferred by NNS vs. NS Spanish academic writers?

RQ3: What specific hedging functions prevail among NNS vs. NS writers? Are certain functions more culturally-specific?

3. Corpus and methods

3.1. Corpus architecture

This contrastive corpus study examines hedging practices across Spanish academic writing by native novice, non-native novice, and native expert writers. Three specialized corpora were compiled following purposive sampling: 21 non-native Spanish master's theses by Chinese students (MT_ZH), 21 native Spanish master's theses (MT_ES), and 34 SSCI-indexed research articles published by native Spanish experts (RA_ES). All collected texts were cleaned to remove non-relevant elements before being manually

formatted and stored for corpus linguistic analysis. Table 1 summarizes the key details of the three corpora analyzed in this study.

Corpus	Number of texts	Genre	L1 status	Year range	Text length (token range)	Total words
MT_ZH	21	Master's thesis	Non-native	(2009-2020)	(12,366-34,434)	424,087
MT_ES	21	Master's thesis	Native	(2009-2020)	(9,846-27,218)	387,613
RA_ES	34	Research article	Native	(2009-2020)	(3,146-8,897)	209,195
Total	**76**					**1,020,895**

Table 1. Description of the three corpora.

To ensure comparability among the three corpora, this study adopted Moreno's (Connor & Moreno, 2005; Moreno, 2008) contrastive model — *tertium comparationis*— which lists a set of comparison criteria (such as text form, mode, tone, discipline, genre, and other dimensions) that help corpus builders to reach corpus comparability judgments. All factors except those of interest need to be constant across the corpora. For instance, cross-modal or cross-disciplinary differences in hedging were not considered in this study, therefore only academic texts in philology-related areas were included in the corpus.

While the two student corpora are fairly comparable, the novice and expert corpora inherently differ in genre because the educational writing of students and professional writing of experts have divergent targeted audience and expertise levels. Though research articles inherently differ in genre from master's theses, these expert-written texts can serve as a useful benchmark representing scholarly writing norms that advanced academic writers strive toward (Ädel, 2006). In this respect, contrasts between the non-native versus native student writing as well as novice versus expert academic writing will still offer valuable implications for teaching Spanish hedging use (cf. Smirnova & Strinyuk, 2020). Finally, multiple comparisons across these three corpora allow a fuller picture of their similarities and differences in hedging behaviors.

3.2. Identification and annotation of hedges

The linguistic forms examined as prospective hedging devices in this study include both Spanish lexical items (i.e., adverbs, adjectives, nouns, and verbs) and one Spanish non-lexical item (i.e., the conditional morpheme *-ría*) (Mur-Dueñas, 2021), as they are typically associated with hedging in the eyes of

analysts (Varttala, 1999). However, no forms are inherently hedgy —a hedge should always be viewed as context-dependent. To systematically verify the hedging function of potential candidates, particularly polysemic and polypragmatic forms, we employed two identification tests following Villalba (2020), Vold (2006), and Crompton (2012): absence test and commutation test. The first one means that the removal of the candidate form should result in the loss of hedging effect, transforming the utterance from hedged to unhedged. Meanwhile, in the second test the candidate form should be replaceable with intrinsic hedging devices (e.g., *es posible que* / 'it is possible that', *a lo mejor* / 'perhaps') without substantial meaning change; or contrastable with intrinsic non-hedging forms (e.g., *es cierto que* / 'it is certain that', *ciertamente* / 'certainly'), resulting in distinctly different degrees of certainty. The application of these tests can be illustrated with the Spanish modal verb *poder*, a typical polypragmatic form. When examining "Puede empezar cuando quiera" ('You can start whenever you want'), neither the absence test (removing "puede" preserves the basic meaning) nor the commutation test (replacement with "es posible que" alters the meaning) indicates hedging function, confirming its role as deontic permission. In contrast, in "Los resultados pueden indicar que..." ('The results could indicate that…'), both tests identify hedging: the absence test shows increased certainty when "pueden" is removed, while the commutation test demonstrates meaning preservation when replaced with the intrinsic hedge *es posible que*, thus confirming its epistemic hedging function.

This study employs a two-phase procedure for hedge annotation in the corpora: initial automatic coding followed by manual verification. In the first phase, textual data was processed with MAXQDA (VERBI Software, 2023) to extract all instances of pre-identified candidate form based on existing hedging inventories (Yao, 2022). This computational extraction provides a broad picture of potential hits. Subsequently in the second phase, the first author manually inspected all extracts in their co-textual environments to filter out false positives, ensuring only genuine hedging cases were retained for analysis. While instances with non-hedging meanings were discarded, those hedges missed during the automatic coding (e.g., *hasta donde sé* / 'to my knowledge', *podemos conjeturar que* / 'we can conjecture that') were incorporated when trawling through the corpora. By leveraging the efficiency of automatic extraction with the accuracy of human validation, this two-step methodology enables a rigorous accounting of hedging behaviors.

All verified hedging devices were finally classified into an eclectic taxonomy that drew on lexico-grammatical and functional principles (Mur-Dueñas, 2021; Prince et al., 1982; Salager-Meyer, 1994; Thomson, 2020; Varttala, 1999). The lexico-grammatical categories this study takes into account are shown in Table 2, and the functional categories, in Table 3. While building on established classifications, our functional categories were adapted to capture the specific patterns observed in our corpus. Notably, we propose a separate category of 'Cautious suggestion' to highlight their distinct rhetorical function in academic writing: expressions of this type serve to strategically frame research implications or recommendations in a less imposing manner.

Category	Examples
Modal verb	**ría* (conditional morpheme), *poder*
Semi-modal verb	*parecer, soler*
Lexical verb	*creer, considerar, sugerir, opinar, asumir*
Adverb	*casi, generalmente, posiblemente, quizá*
Adjective	*posible, probable, habitual*
Noun	*idea, posibilidad, hipótesis*
Phrase	*a menudo, en cierta medida, en general, más o menos*

Table 2. Lexico-grammatical categories of hedges.

Category	Description	Examples
Shield	Reduce writer commitment to a statement	*ser posible, poder, quizás, al parecer*
Approximator	Make a statement less precise	*generalmente, a menudo, casi, normalmente, soler*
Personal doubt	Mark writer's direct involvement in a statement	*en nuestra opinión, creer, considerar*
Cautious suggestion	Politely offer a tentative suggestion	*ser mejor, aconsejar, recomendar*

Table 3. Functional categories of hedges.

To enhance analysis reliability, the second author of this paper and the large language model Claude 3.5 Sonnet (Anthropic, 2024) independently classified 200 instances randomly selected from the corpus, reaching a fairly high inter-annotator agreement based on overlap percentage (94% and 94.5% respectively). The first and second authors later arranged a meeting to discuss the classified segments where the systematic discrepancies occurred in order to improve the codes where necessary and reach a final consensus.

3.3. Analysis of hedges

This study employs a mixed-methods approach to investigate the use of hedging in Spanish academic writing. This integrated approach consists of two main components: a corpus-based quantitative analysis and a complementary qualitative contextual interpretation.

Initially, basic distributional statistics tabulate the frequencies and normalized frequencies of overall hedges as well as specific lexical items and functional categories across the three corpora. Additionally, a one-way ANOVA assesses significant differences between groups, followed by post-hoc Tukey tests pinpointing specific cross-corpus distinctions in hedging densities. Though our data failed the normality assumption (as suggested *a posteriori* by the statistical test software), the one-way ANOVA is considered a robust test against such assumption (Brezina, 2018). This means that it tolerates violations to its normality assumption rather well.

One of the key objectives of this study was to identify the hedging devices that exhibited distinctive usage patterns across the three writer groups. To assess the saliency of these hedges, keyness analysis using log-likelihood (G^2) and Bayesian Information Criterion (BIC) values was conducted (Gabrielatos, 2018). The higher G^2 or BIC scores are, the more salient a certain hedging form is in one corpus compared against the other.

Subsequently, closer manual scrutiny of 500 random instances examines hedge usage in wider co-textual environments. Salient variations in preferred functional and rhetorical hedging strategies are noted through an open inductive coding process. Analysis focuses on uncovering potential linguistic and cultural influences underpinning hedging behaviors by the Chinese students of Spanish compared to native Spanish writers.

Findings synthesize quantitative results signifying broad trends with qualitative insights into contextual motivations, together profiling how conventions in Spanish academic writing manifest in native versus non-native student and expert productions.

4. Results

The quantitative and qualitative findings from the analysis will be now presented. First, the overall frequency of hedges will be outlined to reveal a panorama of hedges distribution across the three corpora. Second, the

frequency distribution of major hedging categories (lexico-grammatical and functional) will be reported, in response to RQ1. Third, the most salient hedging forms will be shown in relation to RQ2. Finally, the specific use of functional categories will be described to answer RQ3.

4.1. Overall frequency of hedging use in the three corpora

The quantitative analysis began by examining the overall distribution of hedging devices employed by each writer group. As shown in Table 4, substantial variations emerged in the density of hedges used across the three corpora.

	MT_ZH	MT_ES	RA_ES
Total number	1,706	2,907	1,213
Per 10,000 words	40.23	74.80	57.98
Frequency mean	81.24	99.86	35.68
Frequency range	21-203	30-167	6-134

Table 4. Frequency of hedges in the three corpora.

The non-native Spanish student writers (MT_ZH) exhibited the lowest overall hedging frequency, with a total of 1,706 instances and a normalized rate of 40.23 hedges per 10,000 words. In contrast, the native Spanish novice writers (MT_ES) demonstrated a considerably higher propensity for hedging, employing 2,907 instances at a normalized frequency of 74.80 per 10,000 words.

Interestingly, the native Spanish expert writers (RA_ES) fell between the two novice groups, using 1,213 hedges with a normalized frequency of 57.98 per 10,000 words. While lower than the native student writers, this rate was still markedly higher than that of the non-native Spanish students.

4.2. Frequency distribution of hedging categories across the corpora

The frequency distributions of major lexico-grammatical hedging categories were analyzed across the three corpora of non-native novice Spanish student writing (MT_ZH), native novice Spanish student writing (MT_ES), and native expert Spanish academic prose (RA_ES). Overall hedging rates displayed some variation, as shown in Table 5 and Figure 1. Modal verbs were the most commonly employed hedge type for all three writer groups.

Category	MT_ZH		MT_ES		RA_ES	
	RF	*NF (pttw)*	*RF*	*NF (pttw)*	*RF*	*NF (pttw)*
Modal verb	316	7.45	843	21.75	503	24.04
Semi-modal verb	213	5.02	312	8.05	158	7.55
Lexical verb	276	6.51	183	4.72	100	4.78
Adverb	416	9.81	304	7.84	205	9.80
Adjective	61	1.44	89	2.30	38	1.82
Noun	17	0.40	42	1.08	16	0.76
Phrase	407	9.60	328	8.46	193	9.23

RF = raw frequency; NF = normalized frequency; pttw = per ten thousand words.

Table 5. Frequency distribution of lexico-grammatical categories across the corpora.

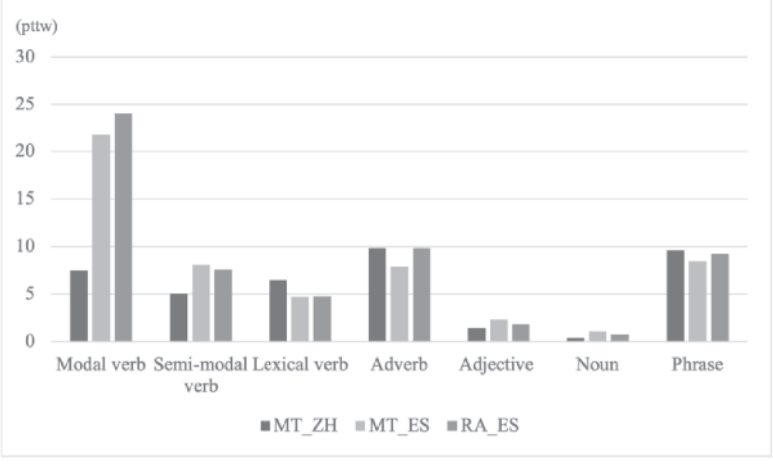

Figure 1. Normalized frequencies of lexico-grammatical categories in the corpora compared.

To assess statistical significance, the one-way ANOVA was conducted, which determined that the mean occurrences of modal verb hedges differed significantly among writer groups ($F(2, 73) = 17.24$, $p < .001$). A Tukey post hoc test revealed that modal verbs employed by the non-native Spanish students were significantly fewer compared to both the native Spanish peers ($p = .018$) and native experts ($p = .007$). However, modal verb frequencies between native novice and native expert groups did not differ significantly ($p = 1.000$). No other lexico-grammatical hedging categories exhibited statistically significant between-group variation based on the one-way ANOVAs.

After examining the lexico-grammatical patterns, this study further analyzed the functional categories of hedges employed by the three writer groups.

Table 6 and Figure 2 present the frequency distributions of major functional hedging categories across the three corpora.

Category	MT_ZH		MT_ES		RA_ES	
	RF	*NF (pttw)*	*RF*	*NF (pttw)*	*RF*	*NF (pttw)*
Shield	575	13.56	1,359	35.06	787	37.62
Approximator	882	20.80	642	16.56	375	17.93
Personal doubt	218	5.14	92	2.37	47	2.27
Cautious suggestion	31	0.73	8	0.21	5	0.23

Table 6. Frequency distribution of functional categories across the corpora.

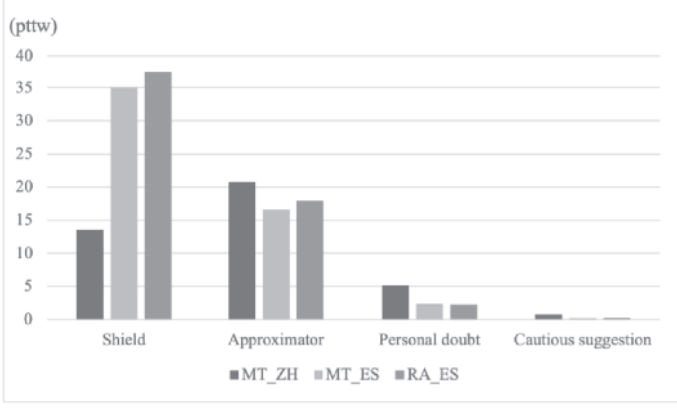

Figure 2. Normalized frequencies of functional categories in the corpora compared.

The category of shields, which express explicit tentativeness or possibility, emerged as the most prevalent functional hedge type for both native writer groups, with the native experts using 787 instances (NF = 37.62 pttw) and the native students employing 1,359 instances (NF = 35.06 pttw). In contrast, the non-native Chinese students relied substantially less on shield hedges, using only 575 instances (NF = 13.56 pttw).

Conversely, the non-native writers demonstrated a stronger preference for approximators, which convey approximation or rounding, employing 882 instances (NF = 20.80 pttw). The native writer groups exhibited relatively lower frequencies of approximator hedges, with 642 instances (NF = 16.56 pttw) for the native students and 375 instances (NF = 17.93 pttw) for the native experts. Regarding personal doubt hedges, which express the writer's

personal uncertainty or lack of confidence, the non-native Chinese students again exhibited the highest frequency.

To assess statistical significance, one-way ANOVA tests were conducted. The results revealed statistically significant differences between the writer groups in their use of shield hedges ($F(2, 73) = 20.14$, $p < .001$). Tukey post hoc tests further showed that for shield hedges, the Chinese students differed significantly from both the native Spanish students ($p = .006$) and the native Spanish experts ($p = .002$), but no significant difference was found between the two native writer groups ($p = 1.000$). In contrast, no statistically significant differences were observed between the three writer groups for the use of other functional categories.

4.3. Comparison of hedging forms across the corpora

The study also tallied the number of distinct forms employed for hedging purposes in each corpus. Intriguingly, results indicated that the Chinese L1 novice writers utilized the greatest variety of distinct hedging forms (106), surpassing both groups of native Spanish writers in the MT_ES (96 forms) and RA_ES (85 forms) corpora.

Furthermore, the frequency pattern of each individual hedging form is reported in order to assess their saliency across the three corpora, as mentioned in Section 3.3. The results presented in Table 7 revealed several noteworthy differences in the deployment of specific hedging forms.

Hedge	MT_ZH		MT_ES		RA_ES		Log-likelihood (G^2)	BIC score
	RF	*NF(pttw)*	*RF*	*NF(pttw)*	*RF*	*NF(pttw)*		
-ría	230	5.42	733	18.91	435	20.79	410.54	382.87
normalmente	112	2.64	16	0.41	7	0.33	98.45	70.78
parecer	75	1.77	192	4.95	89	4.25	68.34	40.67
opinar	36	0.85	0	0.00	0	0.00	63.25	35.58
sugerir	6	0.14	6	0.15	28	1.34	45.43	17.76
generalmente	111	2.62	39	1.01	18	0.86	41.30	13.63
tender	65	1.53	36	0.93	2	0.10	38.97	11.30
a veces	63	1.49	16	0.41	6	0.29	37.71	10.04
probablemente	5	0.12	22	0.57	27	1.29	36.26	8.59
desprender	0	0.00	19	0.49	7	0.33	28.70	1.03

Table 7. The most salient hedging devices across the corpora.

A striking contrast emerged in the use of the conditional morpheme *-ría*, which expresses tentative possibility in Spanish. This hedge was significantly underutilized by the Chinese student writers (RF = 230, NF = 5.42 pttw)

compared to their native Spanish novice (RF = 733, NF = 18.91 pttw) and expert (RF = 435, NF = 20.79 pttw) counterparts.

Another notable finding was the overuse of certain hedging adverbs and verbs by the Chinese students, such as *normalmente* ('normally') (RF = 112, NF = 2.64 pttw), *generalmente* ('generally') (RF = 111, NF = 2.62 pttw), and *tender* (RF = 65, NF = 1.53 pttw), which express generalized claims. These hedges were significantly less frequent in the native Spanish corpora. Furthermore, the Chinese students exhibited a propensity for using the verb *opinar* ('think', 'opine') (RF = 36, NF = 0.85 pttw) as a hedging device, whereas this verb was entirely absent in the native Spanish corpora.

Conversely, the expert writers demonstrated a stronger preference for the hedge *sugerir* (RF = 28, NF = 1.34 pttw) compared to the novice writers. However, the hedging verb *desprender* ('infer') was completely absent in Chinese student writing.

4.4. Specific use of hedging functions across the corpora

While the quantitative patterns revealed distinct preferences for specific functional hedging categories across the three writer groups, a closer qualitative examination uncovers more nuanced variations in how these functions are linguistically realized. By analyzing authentic examples from the corpora, this section aims to provide deeper insights into the rhetorical strategies and pragmatic effects underlying the use of various hedging devices.

4.4.1. Shields

As the quantitative results indicated, the functional category of shield exhibited significant variation across the three writer groups. Specifically, the non-native Chinese students underutilized shield hedges compared to their native Spanish counterparts. Moreover, the lexico-grammatical form that emerged as a particularly salient marker of this underuse was the Spanish conditional morpheme *-ría*.

A closer qualitative examination of authentic examples reveals more nuanced differences in how these writer groups linguistically realize the expression of tentativeness and possibility through shield hedges. The following examples illustrate this variation:

(1) En el 6, la traducción **producirá** cierta confusión sin una clara anotación ni la distinción entre las frases iguales. (MT_ZH_19)

(2) Una extranjerización total, sin adaptaciones o ampliaciones de ningún tipo, **produciría** incomprensión y extrañamiento en el lector, abrumado por una realidad que no es la suya. (MT_ES_21)

While the non-native writer states "la traducción producirá cierta confusión" ('the translation will produce some confusion'), the native writer more tentatively hedges using the conditional "produciría incomprensión y extrañamiento" ('it would produce incomprehension and estrangement'). The conditional form *-ría* effectively conveys a more tentative and less definitive stance, aligning with disciplinary conventions of hedging in Spanish academic writing.

(3) De los estudios pragmáticos de la atenuación **destacan**, entre otras, las obras de Fraser (1980), Holmes (1984), Haverkate (1992), Caffi (1999) y Sbisá (2001), las cuales se caracterizan por su dependencia de las teorías pragmáticas como la de los actos de habla y la de la cortesía. (MT_ZH_07)

(4) En la última década son muchos los estudios encaminados a esclarecer o categorizar el papel del intérprete en el ámbito sanitario. **Destacarían** modelos como el de abogacía e imparcialidad propuesto por Cambridge (2002) o el enfoque de Leanza (2005) sobre el papel del intérprete como facilitador de la comunicación o asimilador cultural (RA_ES_29)

Similarly, the non-native writer presents information as relatively direct facts: "De los estudios pragmáticos [...] destacan [...] las obras de [...]" ('From pragmatic studies stand out the works of […]'). In contrast, the native expert exemplifies more skilled hedging by using "Destacarían modelos como [...]" ('Models such as […] would stand out'), employing the conditional *-rían* to express tentativeness and leave room for alternative perspectives.

These contrastive examples showcase the native writers' pragmatic competence in deploying the conditional mood to express possibilities and attenuate the certainty of their claims, a strategy that is notably underutilized by the non-native writers. Such pragmatic nuances in using linguistic devices like the conditional to realize the shield function are crucial for appropriate hedging in Spanish academic discourse.

4.4.2. Approximators

The quantitative analysis revealed no statistically significant differences in the overall use of approximator hedges across the three writer groups. However,

certain specific approximator devices like *normalmente*, *generalmente*, and *tender* ('tend') emerged as relatively overused by the non-native Chinese student writers compared to their native Spanish counterparts.

To better understand the pragmatic nuances underlying this quantitative pattern, a qualitative examination of authentic examples focused on the salient approximator *normalmente*. This analysis revealed some notable variations in how the different writer groups linguistically realized and rhetorically operationalized this hedging form.

(5) …por los diferentes sistemas de escritura, el préstamo en la traducción entre el chino y el español se refiere **normalmente** a la transliteración (MT_ZH_20)

(6) El aprendizaje situado preconiza que el conocimiento se ha de presentar en el contexto auténtico en que **normalmente** se desenvuelve. (RA_ES_33)

The non-native writers exhibited a tendency to use *normalmente* in more generalized contexts, making broad approximations about behaviors or conceptual associations. For instance, "normalmente" in Example 5 overgeneralizes the referential meaning of a terminology. In contrast, native writers, especially experts, deployed "normalmente" for more specific, discipline-relevant approximations, as in Example 6.

(7) Por las diferencias culturales, **normalmente** los términos con significado figurado no se pueden traducir literalmente, a menos que se añadan explicaciones extras. (MT_ZH_20)

(8) **Normalmente** son las dificultades comunicativas, sobre todo en cuanto a producción del discurso, las que alertan al hablante del proceso de erosión que ha sufrido una lengua. (MT_ES_01)

Additionally, non-native productions sometimes reflected a potentially excessive or unnecessary use of *normalmente* in contexts where hedging seemed pragmatically unwarranted, because it is a well-established given, as exemplified by Extract 7. Native writers, on the other hand, appeared more judicious in their contextually appropriate deployment of this approximator hedge, as in Extract 8.

4.4.3. Personal doubts

While the quantitative results did not reveal statistically significant differences in the overall use of 'personal doubt' hedges across the three writer groups, a qualitative analysis unveiled interesting variations in how writers linguistically manifested their presence and expressed personal uncertainty or lack of confidence in their claims.

The non-native student writers exhibited a propensity for overtly involving themselves when presenting arguments or opinions. This was evident in their frequent use of first-person plural pronouns (e.g., *creemos* / 'we think', *consideramos* / 'we consider', *opinamos* / 'we think') and first-person plural possessives (e.g., *en nuestra opinión* / 'in our opinion'). For instance:

(9) Entonces, **opinamos** que la Comprensión de Lectura de EEE8 ha cumplido con el requisito pertinente de una manera satisfactoria. (MT_ZH_06)

(10) En **nuestra** opinión, esto coincide justamente con los problemas que hay que tener en cuenta a la hora de traducir los culturemas. (MT_ZH_20)

In contrast, the native Spanish writers, particularly the expert authors, tended to maintain a more detached authorial presence, less conspicuously marking their personal doubt or uncertainty (e.g., *se considera* / 'it is considered', *se cree* / 'it is thought'). However, when native writers (novice especially) did express personal stance, they sometimes used first-person singular pronouns (e.g., *considero* / 'I consider') and first-person singular possessives (e.g., *a mi juicio* / 'in my opinion'), as exemplified in Extract 11. Notably, non-native student writers also employed first-person singular forms to express personal doubt, but they predominantly relied on the informal phrase *creo que* ('I think that'), which was entirely absent from the native writer corpora (see Extract 12).

(11) **Considero** que todas estas propiedades del fenómeno pueden ser explicadas de manera simple en base al siguiente principio: (MT_ES_08)

(12) **Creo** que aquí la traducción literal más nota resultará mejor. (MT_ZH_19)

4.4.4. Cautious suggestion

While the quantitative data did not reveal statistically significant differences in the overall use of 'cautious suggestion' hedges across the three writer groups, likely due to the low frequency of such cases in the corpora (see Table 6), a qualitative analysis provided valuable insights into how the three writer groups linguistically realized suggestions or recommendations.

One notable finding was the native writers' preference for using the conditional mood (*-ría* form) to hedge and mitigate the imposition of their suggestions, as exemplified in Extract 13. In contrast, the L2 Spanish student writers tended to employ more direct formulations when making suggestions, often lacking the pragmatic nuance of mitigation through conditional forms, as seen in Excerpt 14. This finding aligns with previous observations regarding the non-native writers' underuse of the conditional morpheme *-ría* as an effective hedging device for expressing tentativeness and possibility (see Section 4.4.1).

(13) **Sería** conveniente, además, que las actividades de lectura fueran colectivas e interactivas, y que se trabajara con personas de distintas nacionalidades. (MT_ES_06)

(14) Por lo tanto, para que el lector pueda comprender más o menos lo mismo que el lector del texto origen, opinamos que **será** conveniente añadir una nota al pie de página explicando este "chengyu" y su relación de esta frase con el fin de conseguir una mejor equivalencia funcional. (MT_ZH_20)

Additionally, the non-native student writers appear to more overtly involve themselves when making suggestions or recommendations, frequently using first-person plural pronouns, as in Extract 15. This pattern also resonates with the non-native writers' tendency to overtly mark their authorial presence through plural first-person forms, as observed in the previous section.

(15) Entonces **recomendamos** hacer pruebas también en otros países. (MT_ZH_05)

5. Discussion

This study investigated hedging practices across Spanish academic writing by non-native novice, native novice, and native expert writers. The quantitative

analyses revealed distinct patterns in the deployment of various lexico-grammatical and functional hedging categories. In general, hedges were employed less frequently in L1 Chinese student texts than native Spanish texts. This finding is in line with previous studies which point out that non-native writers may not have sufficient command of linguistic resources to express hedging (Vassileva, 2001; Yang, 2013). Notably, the non-native Chinese student writers underutilized certain salient forms like the conditional morpheme *-ría* for expressing tentativeness, while overusing some approximator adverbs like *normalmente* and *generalmente*.

The qualitative examination of authentic examples further elucidated pragmatic variations underlying these quantitative trends. The native writers, especially experts, demonstrated a nuanced, contextually-appropriate use of hedges aligning with disciplinary conventions. In contrast, the non-native productions sometimes exhibited overgeneralized or excessive hedging tendencies deviating from pragmatic norms.

With this overview of the key findings, the following subsections discuss potential linguistic, cultural, and expertise factors influencing the observed hedging behaviors across writer groups.

5.1. Linguistic factor

One of the most striking findings was the non-native Chinese student writers' significant underuse of the conditional morpheme *-ría* for expressing tentativeness compared to their native Spanish counterparts. Several linguistic factors could potentially contribute to this variation.

Firstly, *-ría* is a non-lexical item, a suffix that is less perceptually prominent in comparison with other individual lexical items. This morpheme also bears a specific pragmatic meaning of expressing epistemic possibility, a concept arguably absent from Chinese. As Chinese does not use morphological inflections to convey complex hypothetical or uncertain meaning, the Spanish conditional tense, marked systematically with the suffix *-ría*, poses an entirely unfamiliar and challenging structure for L1 transfer or analogy. Chinese L1 writers must learn to comprehend and produce conditional forms like *-ría* as largely separate lexical items, memorizing appropriate context usage on a case-by-case basis. This adds substantial cognitive load for developing pragmalinguistic skills. In contrast, common hedging constructions like *poder* + infinitive (Chao Parapar, 2018) or other lexical hedges like *ser posible* and *quizás* ('perhaps')

may be simpler for Chinese learners to grasp without needing to process inflections.

5.2. Cultural factor

Beyond linguistic factors, the observed variations in how writers linguistically manifested their authorial presence and expressed personal stance appear to be shaped by cultural norms and rhetorical traditions. Findings in the earlier sections indicate that the Chinese student writers are inclined to overtly involve themselves through frequent use of first-person plural pronouns and possessives (e.g., *creemos*, *opinamos*, *a nuestro entender* / 'in our opinion', *aconsejamos* / 'we suggest') when hedging arguments or making suggestions. This tendency among the non-native writers can be attributed to the collective culture in Chinese academic writing (Chen & Zhang, 2017; Wang & Jiang, 2018), where an overtly plural authorial voice is more conventionally accepted, even in single-authored texts. It could be argued here that such plural self-mentions may serve to create a sense of solidarity and diffuse individual responsibility, reflecting cultural values around collectivism and modesty norms.

In contrast, the native Spanish writers, especially experts, tended to maintain a more detached authorial presence and favor a more impersonal or depersonalized style, a common hedging strategy (Luukka & Markkanen, 1997) when stating opinions. This preference is aligned with the general academic writing traditions where the objectivity and detachment are foregrounded while the personal involvement and explicit authorial presence are discouraged (Chafe, 1982; Harwood, 2005; Uba & Baynham, 2017).

Interestingly, while the non-native writers exhibited a strong preference for plural self-mentions, the use of singular first-person forms was almost entirely absent from their writing, roughly in line with Wang and Jiang's (2018) finding on the self-mention use between Chinese student writers and English expert writers. In the case of native Spanish writers, however, they do occasionally use singular self-mentions (e.g., *considero*, *a mi juicio*) to present personal opinions. This contrast could stem from different institutional teaching culture (Yao, 2022), where Chinese writers are taught not to present themselves individually but collectively (Hyland, 2002; Li & Wharton, 2012).

5.3. L2 expertise/proficiency factor

Our analysis also revealed that several unconventional or atypical hedging devices appeared exclusively in the non-native corpus, which could explain why Chinese L1 novice writers employed more diverse hedging forms than other two writer groups, as mentioned in Section 4.3. Expressions such as *adivinar* ('guess'), *opinar*, *deber de* ('should'), *creo que*, *nos parece que* ('we think that'), *a lo mejor*, and *por lo visto* ('apparently') were found only in the Chinese students' texts, while being absent from both native novice and expert corpora. On the other hand, advanced words like *sugerir* ('hint, suggest') and *desprender* were either rarely used or not used in L2 texts, suggesting that L2 expertise and proficiency play a crucial role in the appropriate deployment of hedges in academic writing.

This phenomenon can be attributed to several factors related to L2 proficiency and pragmatic competence. Firstly, as Hyland and Milton (1997) and Hinkel (2005) show, L2 learners often struggle with the nuanced deployment of epistemic devices, leading to an overextension of their hedging repertoire. This overextension may stem from their limited academic writing skills, which make them use more familiar hedging resources (Smirnova & Strinyuk, 2020).

Moreover, the use of informal and colloquial phrases like *creo que*, *nos parece que*, *a lo mejor*, and *por lo visto* in formal academic prose indicates a lack of register awareness among non-native writers. This result confirms Hinkel (2005) and Lee et al.'s (2019) finding that informal features generally appear more often in L2 texts. It is perhaps understandable as developing an understanding of appropriate academic register is a challenging aspect of L2 writing, often requiring explicit instruction and extensive exposure to disciplinary discourse.

Fortunately, from a developmental perspective, advanced learners gradually acquire more sophisticated and context-appropriate rhetorical strategies (Carrió-Pastor, 2021). That is, as L2 proficiency and academic writing expertise increase, learners develop a more nuanced understanding of appropriate hedging strategies.

6. Conclusion

This study examined hedging practices across Spanish academic writing by non-native novice, native novice, and native expert writers. The findings

address three key research questions, shedding light on the complex interplay of linguistic, cultural, and expertise factors in academic hedging.

Regarding hedging categories, both non-native and native writers employed a range of lexico-grammatical and functional devices. However, significant differences emerged in the use of modal verb hedges, with non-native writers using significantly fewer compared to both native novice and expert writers. Similarly, for functional categories, shield hedges showed significant variation, with non-native writers differing significantly from both native groups. Individual hedging form preferences also varied notably. Non-native writers exhibited a salient overuse of certain approximators like *normalmente* and *generalmente*, while significantly underutilizing the conditional morpheme *-ría*. In contrast, native writers demonstrated a more diverse hedging repertoire, with expert deployment of the conditional form. Finally, concerning hedging functions, all groups utilized shields, approximators, personal doubt markers, and cautious suggestions. However, non-native writers showed a distinct tendency towards more direct and over-generalized expressions, as well as overt self-mention through first-person plural forms and informal expressions, contrasting with the more impersonal or nuanced stance-taking of native writers, who often employed more detached and context-appropriate hedging forms when expressing personal stance.

These variations reflect linguistic challenges, cultural influences, and L2 expertise factors. The absence of equivalent conditional structures in Chinese likely contributes to the underuse of *-ría* by non-native writers. Similarly, the preference for plural self-mention among non-native writers aligns with Chinese academic conventions emphasizing collective voice, contrasting with the more individualistic or impersonal Spanish norms. Furthermore, the exclusive appearance of unconventional hedging devices (e.g., *opinar*, *creo que*, *a lo mejor*) in non-native texts suggests that L2 proficiency plays a crucial role in the appropriate deployment of hedges.

The findings have important pedagogical implications. Explicit instruction in the pragmatic functions of hedging devices, particularly the conditional mood, could benefit L2 writers. Raising awareness of cross-cultural differences in authorial presence and providing exposure to discipline-specific hedging norms could enhance L2 learners' academic writing competence (Vold, 2006). Moreover, focused attention on register awareness could help non-native writers avoid inappropriate colloquial hedges in formal academic prose (Hinkel, 2005).

While this study offers valuable insights, it is not without limitations. The lack of interview data from the writers themselves restricts our understanding of their conscious hedging choices and motivations. Future research could address this by incorporating writer interviews (see for example Chen & Zhang, 2017) to gain deeper insights into the cognitive processes and cultural influences underlying hedging practices. Additionally, longitudinal studies tracking the development of hedging competence among L2 writers and incorporating non-native expert writing in our corpus architecture could further illuminate the interplay between language proficiency, academic socialization, and hedging deployment.

Article history:
Received 7 August 2024
Received in revised form 5 December 2024
Accepted 30 December 2024

References

Abdi, R. (2002). Interpersonal metadiscourse: An indicator of interaction and identity. *Discourse Studies, 4*(2), 139-145. https://doi.org/10/c6j3mh

Abdollahzadeh, E. (2019). A cross-cultural study of hedging in discussion sections by junior and senior academic writers. *Ibérica, Journal of the European Association of Languages for Specific Purposes, 38*, 177-202. https://revistaiberica.org/index.php/iberica/article/view/97

Acín Villa, E. (2016). La atenuación en las "Conclusiones" de las tesis de doctorado. *Textos En Proceso, 2*(1), 1-24.

Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. John Benjamins.

Akbas, E. (2012). Exploring metadiscourse in Master's dissertation abstracts: Cultural and linguistic variations across postgraduate writers. *International Journal of Applied Linguistics & English Literature, 1*(1), 12-26. https://doi.org/10/gg6bs5

Albelda, M. (2018). La variación genérico-discursiva de la atenuación como resultado de la variación de la imagen. *Spanish in Context, 15*(2), 346-368. https://doi.org/10.1075/sic.00018.alb

Albelda, M. (2020). On the mitigating function of the Spanish evidential *se ve que*. *Corpus Pragmatics, 4*(1), 83-106. https://doi.org/10.1007/s41701-019-00067-8

Albelda, M., & Estellés, M. (2021). Mitigation revisited. An operative and integrated definition of the pragmatic concept, its strategic values, and its linguistic expression. *Journal of Pragmatics, 183*, 71-86. https://doi.org/10.1016/j.pragma.2021.07.002

Albelda, M., Briz, A., Cestero, A. M., Kotwica, D., & Villalba, C. (2014). Ficha metodológica para el análisis pragmático de la atenuación en corpus discursivos del español (ES.POR.ATENUACIÓN). *Oralia: Análisis del Discurso Oral, 17*, 7-62. https://doi.org/10.25115/oralia.v17i.7999

Alonso Alonso, R. (2019). A multicompetence perspective of hedging in second language academic writing. *Onomázein*, *45*, 58-78. https://doi.org/10.7764/onomazein.45.02

Anthropic. (2024). *Claude 3.5 Sonnet (June 20 version)* [Large language model]. https://claude.ai/

Aull, L. L., & Lancaster, Z. (2014). Linguistic markers of stance in early and advanced academic writing: A corpus-based comparison. *Written Communication, 31*(2), 151-183. https://doi.org/10/f57m3n

Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press.

Briz, A. (1995). La atenuación en la conversación coloquial: Una categoría pragmática. In L. Cortés (Ed.), *El español coloquial: Actas del I Simposio sobre análisis del discurso oral* (pp. 103-122). Universidad de Almería.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.

Caffi, C. (1999). On mitigation. *Journal of Pragmatics, 31*(7), 881-909. https://doi.org/10.1016/S0378-2166(98)00098-8

Cañada Pujols, M. D., & Bach, C. (2022). Marcadores metadiscursivos interaccionales en resúmenes de TFG: ¿indicadores de dominio del género? *Círculo de Lingüística Aplicada a la Comunicación, 90,* 5-19. https://doi.org/10.5209/clac.81301

Carrió-Pastor, M. L. (2016). A contrastive study of the hedges used by English, Spanish and Chinese researchers in academic papers. In F. Alonso Almeida, I. Ortega Barrera, E. Quintana Toledo, & M. E. Sánchez Cuervo (Eds.), *Input a word, analyze the world: Selected approaches to corpus linguistics* (pp. 477-491). Cambridge Scholars.

Carrió-Pastor, M. L. (2019). Phraseology in specialised language: A contrastive analysis of mitigation in academic papers. In G. Corpas Pastor & R. Mitkov (Eds.), *Computational and corpus-based phraseology* (pp. 61-72). Springer.

Carrió-Pastor, M. L. (2021). The assessment of metadiscourse devices in English as a foreign language. *Assessing Writing*, *50*. https://doi.org/10.1016/j.asw.2021.100560

Cestero, A. M., & Albelda, M. (2020). Estudio de variación en el uso de atenuación I: Hacia una descripción de patrones dialectales y sociolectales de la atenuación en español. *Revista Signos, 53*(104), 935-961. https://doi.org/10.4067/S0718-09342020000300935

Chafe, W. (1982). Integration and involvement in speaking, writing and oral literature. In D. Tannen (ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35-53). Ablex.

Channell, J. (1985). Vagueness as a conversational strategy. *Nottingham Linguistic Circular*, *14*, 3-24.

Channell, J. (1994). *Vague language.* Oxford University Press.

Chao Parapar, M. (2018). Procedimientos de atenuación en los artículos de investigación científica: Las directrices cognitivas en español. In A. García Ramón & M. A. Soler Bonafont (Eds.), *Estudios de atenuación en el discurso* (pp. 139-156). ELUA, Anexo IV. https://doi.org/10.14198/ELUA2018.Anexo4.08

Chen, C., & Zhang, L. J. (2017). An intercultural analysis of the use of hedging by Chinese and Anglophone academic English writers. *Applied Linguistics Review, 8*(1), 1-34. https://doi.org/10.1515/applirev-2016-2009

Clyne, M. (1994). *Inter-cultural communication at work: Cultural values in discourse*. Cambridge University Press.

Connor, U., & Moreno, A. I. (2005). Tertium comparationis: A vital component in contrastive rhetoric research. In P. Bruthiaux, D. Atkinson, W. G. Eggington, W. Grabe & V. Ramanathan (Eds.), *Directions in applied linguistics: Essays in honor of Robert B. Kaplan* (pp. 153-164). Multilingual Matters.

Crompton, P. (1997). Hedging in academic writing: Some theoretical problems. *English for Specific Purposes, 16*(4), 271-287. https://doi.org/10.1016/S0889-4906(97)00007-0

Crompton, P. (2012). Characterising hedging in undergraduate essays by Middle Eastern students. *The Asian ESP Journal, 8*(2), 55-78.

Dontcheva-Navratilova, O. (2016). Cross-cultural variation in the use of hedges and boosters in academic discourse. *Prague Journal of English Studies, 5*(1), 163-184. https://doi.org/10.1515/pjes-2016-0009

Dudley-Evans, T. (2000). Genre analysis: A key to a theory of ESP? *Ibérica, Journal of the European Association of Languages for Specific Purposes, 2*, 3-11.

Figueras, C. (2018). Atenuación, género discursivo e imagen. *Spanish in Context, 15*(2), 258-280. https://doi.org/10.1075/sic.00014.fig

Fraser, B. (1980). Conversational mitigation. *Journal of Pragmatics, 4*(4), 341-350. https://doi.org/10.1016/0378-2166(80)90029-6

Fraser, B. (2010). Pragmatic competence: The case of hedging. In G. Kaltenböck, W. Mihatsch & S. Schneider (Eds.), *New approaches to hedging* (pp. 15-34). Emerald.

Gabrielatos, C. (2018). Keyness analysis: Nature, metrics and techniques. In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse: A critical review* (pp. 225-258). Routledge.

Harwood, N. (2005). 'Nowhere has anyone attempted … In this article I aim to do just that': A corpus-based study of self-promotional *I* and *we* in academic writing across four disciplines. *Journal of Pragmatics, 37*(8), 1207-1231. https://doi.org/10.1016/j.pragma.2005.01.012

Hinkel, E. (2005). Hedging, inflating, and persuading in L2 academic writing. *Applied Language Learning, 15*(1&2), 29-53.

Hu, G., & Cao, F. (2011). Hedging and boosting in

abstracts of applied linguistics articles: A comparative study of English- and Chinese-medium journals. *Journal of Pragmatics, 43*(11), 2795-2809. https://doi.org/10.1016/j.pragma.2011.04.007

Hyland, K. (1996a). Talking to the academy: Forms of hedging in science research articles. *Written Communication, 13*(2), 251-281. https://doi.org/10.1177/0741088396013002004

Hyland, K. (1996b). Writing without conviction? Hedging in science research articles. *Applied Linguistics, 17*(4), 433-454. https://doi.org/10.1093/applin/17.4.433

Hyland, K. (1998). *Hedging in scientific research articles*. John Benjamins.

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics, 34*(8), 1091-1112. https://doi.org/10.1016/S0378-2166(02)00035-8

Hyland, K. (2005a). *Metadiscourse: Exploring interaction in writing*. Continuum.

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., & Jiang, F. K. (2016). Change of attitude? A diachronic study of stance. *Written Communication, 33*(3), 251-274. https://doi.org/10.1177/0741088316650399

Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing, 6*(2), 183-205. https://doi.org/10.1016/S1060-3743(97)90033-3

Lakoff, G. (1973). Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic, 2*, 458-508.

Lee, J. J., Bychkovska, T., & Maxwell, J. D. (2019). Breaking the rules? A corpus-based comparison of informal features in L1 and L2 undergraduate student writing. *System, 80*, 143-153. https://doi.org/10.1016/j.system.2018.11.010

Li, T., & Wharton, S. (2012). Metadiscourse repertoire of L1 Mandarin undergraduates writing in English: A cross-contextual, cross-disciplinary study. *Journal of English for Academic Purposes, 11*(4), 345-356. https://doi.org/10.1016/j.jeap.2012.07.004

Liu, C., & Tseng, M.-Y. (2021). Paradigmatic variation in hedging and boosting: A comparative study of discussions in narrative inquiry and grounded theory research. *English for Specific Purposes, 61*, 1-16. https://doi.org/10.1016/j.esp.2020.08.002

Luukka, M. R., & Markkanen, R. (1997). Impersonalization as a form of hedging. In R. Markkanen & H. Schröder (Eds.), *Hedging and discourse: Approaches to the analysis of a pragmatic phenomenon in academic texts* (pp. 168-187). De Gruyter. https://doi.org/10.1515/9783110807332.168

Markkanen, R., & Schröder, H. (1997). Hedging: A challenge for pragmatics and discourse analysis. In R. Markkanen & H. Schröder (Eds.), *Hedging and discourse: approaches to the analysis of a pragmatic phenomenon in academic texts* (pp. 3-18). De Gruyter. https://doi.org/10.1515/9783110807332.3

Martín-Martín, P. (2008). The mitigation of scientific claims in research papers: A comparative study. *International Journal of English Studies, 8*(2), 133-152.

Meyer, P. G. (1997). Hedging strategies in written academic discourse: Strengthening the argument by weakening the claim. In R. Markkanen & H. Schröder (Eds.), *Hedging and discourse: approaches to the analysis of a pragmatic phenomenon in academic texts* (pp. 21-41). De Gruyter.

Moreno, A. I. (2008). The importance of comparable corpora in cross-cultural studies. In U. Connor, E. Nagelhout & W. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 25-41). John Benjamins.

Mur-Dueñas, P. (2021). There may be differences: Analysing the use of hedges in English and Spanish research articles. *Lingua, 260*, 103131. https://doi.org/10.1016/j.lingua.2021.103131

Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics, 10*(1), 1-35. https://doi.org/10.1093/applin/10.1.1

Navarro, F., Montes, S., & Álvarez, M. (2022). How do students write in engineering and the humanities? Intertextuality and metadiscourse in undergraduate dissertations written in Spanish. *Círculo de Lingüística Aplicada a la Comunicación, 90*, 35-46. https://doi.org/10.5209/clac.81305

Núñez-Román, F., Hunt-Gómez, C., Moreno-Crespo, P., & Ballesteros-Moscosio, M. Á. (2021). Género y voz del autor en textos académicos de estudiantes de Ciencias de la Educación. In R. G. Muñoz, H. Gonçalves Pinto, I. S. Dias, M. O. Abreu, & D. Alves (Eds.), *Investigação, Práticas e Contextos em Educação 2021* (pp. 66-73). Instituto Politécnico de Leiria.

Poole, R., Gnann, A., & Hahn-Powell, G. (2019). Epistemic stance and the construction of knowledge in science writing: A diachronic corpus study. *Journal of English for Academic Purposes*,

42, 1-11. https://doi.org/10.1016/j.jeap.2019.100784

Prince, E. F., Frader, J., & Bosk, C. (1982). On hedging in physician-physician discourse. In R. J. di Pietro (Ed.), *Linguistics and the professions* (pp. 83-97). Ablex.

Salager-Meyer, F. (1994). Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes, 13*(2), 149-170. https://doi.org/10.1016/0889-4906(94)90013-2

Salager-Meyer, F. (1997). I think that perhaps you should: A study of hedges in written scientific discourse. In T. Miller (Ed.), *Functional approaches to written text: Classroom applications* (pp. 105-118). English Language Programs.

Sanjaya, I. N. S. (2015). Revisiting the effects of sociocultural context and disciplines on the use of hedges in research articles. *Indonesian Journal of Applied Linguistics, 5*(1), 116-125. https://doi.org/10.17509/ijal.v5i1.838

Schröder, H., & Zimmer, D. (1997). Hedging research in pragmatics: A bibliographical research guide to hedging. In R. Markkanen & H. Schröder (Eds.), *Hedging and discourse: approaches to the analysis of a pragmatic phenomenon in academic texts* (pp. 249-272). De Gruyter. https://doi.org/10.1515/9783110807332.249

Smirnova, E., & Strinyuk, S. (2020). Hedges in Russian EAP writing: A corpus-based study of research papers in management. *Journal of English as a Lingua Franca, 9*(1), 81-101. https://doi.org/10.1515/jelf-2020-2033

Thomson, J. J. (2020). Metadiscourse in upper secondary pupil essays: Adapting a taxonomy. *Nordic Journal of Literacy Research, 6*(1), 26-48. https://doi.org/10.23865/njlr.v6.1720

Uba, S. Y., & Baynham, M. (2017). Constraints on authorial stance in accounting PhD theses in a Nigerian university. In Ç. Hatipoğlu, E. Akbas & Y. Bayyurt (Eds.), *Metadiscourse in written genres: Uncovering textual and interactional aspects of texts* (pp. 109-138). Peter Lang.

Varttala, T. (1999). Remarks on the communicative functions of hedging in popular scientific and specialist research articles on medicine. *English for Specific Purposes, 18*(2), 177-200. https://doi.org/10.1016/S0889-4906(98)00007-6

Varttala, T. (2001). *Hedging in scientifically oriented discourse: Exploring variation according to discipline and intended audience*. Doctoral dissertation. University of Tampere.

Vassileva, I. (1997). Hedging in English and Bulgarian academic writing. In A. Duszak (Ed.), *Culture and styles of academic discourse* (pp. 203-222). De Gruyter. https://doi.org/10.1515/9783110821048.203

Vassileva, I. (2001). Commitment and detachment in English and Bulgarian academic writing. *English for Specific Purposes, 20*(1), 83-102. https://doi.org/10.1016/S0889-4906(99)00029-0

VERBI Software. (2023). *MAXQDA 2024* [Computer software]. www.maxqda.com

Villalba, C. (2020). Recognising mitigation: Three tests for its identification. *Journal of Pragmatics, 167*, 68-79. https://doi.org/10.1016/j.pragma.2020.06.015

Vold, E. T. (2006). Epistemic modality markers in research articles: A cross-linguistic and cross-disciplinary study. *International Journal of Applied Linguistics, 16*(1), 61-87. https://doi.org/10.1111/j.1473-4192.2006.00106.x

Wang, J., & Jiang, F. K. (2018). Epistemic stance and authorial presence in scientific research writing: Hedges, boosters and self-mentions across disciplines and writer groups. In P. Mur-Dueñas & J. Šinkūnienė (Eds.), *Intercultural perspectives on research writing* (pp. 195-216). John Benjamins.

Yang, Y. (2013). Exploring linguistic and cultural variations in the use of hedges in English and Chinese scientific discourse. *Journal of Pragmatics, 50*(1), 23-36. https://doi.org/10.1016/j.pragma.2013.01.008

Yao, G. (2022). *Metadiscourse use in Spanish academic writing: Exploring the interface of nativeness and expertise*. Doctoral dissertation. Universidad de Murcia.

Yao, M., Wei, Y., & Wang, H. (2023). Promoting research by reducing uncertainty in academic writing: A large-scale diachronic case study on hedging in Science research articles across 25 years. *Scientometrics, 128*(8), 4541-4558. https://doi.org/10.1007/s11192-023-04759-6

Gang Yao is a Lecturer in the Department of Hispanic and Portuguese Studies at Beijing Foreign Studies University (China). His main research interests are within corpus linguistics, academic discourse, and Spanish as a

Foreign Language. His publications have appeared in journals such as *Ibérica* and *Research in Corpus Linguistics*.

Wei Sun (corresponding author) is a Lecturer in the Department of Hispanic and Portuguese Studies at Beijing Foreign Studies University (China). His main research interests include translation studies and corpus linguistics. He has published in journals such as *Onomázein*, *ALPHA*, and *Revista de Lingüística Teórica y Aplicada*.

NOTES

Detailed supplementary materials of this study such as the articles and thesis collected for the corpora and the prompt for AI annotation have been stored on OSF (https://osf.io/rntpm/?view_only=d357fc9de1094167a7065c69e5fee3fc).