811.111'367.52
https://doi.org/10.18485/bells.2022.14.2

Elisa Ghia[*]
University of Pavia

Emilia Petrocelli[**]
University for Foreigners of Siena

Sergio Pizziconi[***]
University for Foreigners of Siena

DATA POSSIBLY SUGGEST... HEDGING IN SECOND LANGUAGE WRITING: A STUDY ON ADVANCED ITALIAN EFL LEARNERS

Abstract
This paper explores the use of hedges in academic thesis abstracts among L1-Italian advanced university EFL learners drawing on a corpus of 217 abstracts written in English. Two rounds of contrastive analysis are carried out: 15 hedges are first compared to boosters and modals of logical possibility and inferential certainty. Secondly, a comparison is drawn against the MICUSP Corpus.

Results show that EFL students tend to use fewer hedges than their L1-English peers in favor of boosters and modals of epistemic certainty. However, patterns in the two corpora follow certain parallelisms and suggest awareness of hedging strategies by EFL learners. Given its complexity as a linguistic phenomenon, hedging deserves greater attention in academic writing courses and textbooks.

[*] elisa.ghia@unipv.it
[**] petrocelli@unistrasi.it
[***] sergio.pizziconi@unical.it

1. Hedging as a metadiscursive practice

Academic writing is generally typified by a distinctive text structure and organization, the use of specialized lexicon, and the recurrence of typical syntactic patterns, such as complex subordination and passive-voice sentences. Another prominent feature of academic discourse is hedging, a metadiscursive phenomenon at the boundary between lexicon, syntax, and pragmatics. Through hedges, writers establish an interpersonal relationship with their readers and the scientific community at large. Hedge words are an expression of interpersonal stance and add caution and accuracy to statements by lowering the assertive force of a claim and its universal generalizability, thus avoiding "categorical assertions of truth" (Hyland 1998: 6). Hedging is characteristic of science discourse as it makes statements more reliable and widely acceptable, and enables writers to negotiate knowledge with their readers. In these terms, hedging facilitates recognition in the academic community and is also a component of argumentative and persuasive discourse. It amplifies credibility by modulating levels of assertiveness and trustworthiness and by avoiding sharp tendencies in favor of balanced propositions to argue claims.

In a broad view, hedge words include all those linguistic items that carry an epistemic stance and express the degree of truth of a proposition, limitations, imprecision, and approximation (Biber et al. 1999: 557). In addition, hedges also carry affective meaning, in that they show the writer's attitude towards readers and the target subject matter (Hyland 1998).

In the vast literature on interpersonal stance, hedges have been variously classified. Common taxonomies are based on their scope, their specific function, and, most typically, the morphosyntactic category they belong to (their lexico-grammatical and surface features, according to Hyland 1998).

In terms of scope, hedges can lower the assertiveness or degree of certainty of a single word, a phrase, a full sentence, or even the illocutionary force of an entire utterance (Fraser 2010: 203-4). Hedging devices may more specifically be used to attenuate a statement and minimize a threat to the writer's face, thus conveying "evasiveness, tentativeness, fuzziness, mitigation of responsibility and/or mitigation of certainty to the truth value of a proposition" (Salager-Meyer 1997: 129; Biber et al. 1999); conversely, they can also display greater accuracy of the author by revealing "the true state of the writers' understanding and may be used to

negotiate an accurate representation of the state of the knowledge under discussion" (Salager-Meyer, 1997: 129; Hyland 1994; Garzone 2006). Overall, hedges can express politeness by mitigating an assertion while concurrently acknowledging other scholars' work and stance, and allow for acceptability within the scientific community while aligning with the generally recognized standards of academic writing (Hinkel 2005).

Syntactically speaking, hedge words in written English include items from different word classes, ranging from adverbs to verbs and multiword units. Some of the most represented categories are epistemic stance adverbs (e.g., *possibly*, *potentially*, *apparently*), approximating adverbs and downtoners (e.g., *quite, almost*, *virtually*), modal verbs (e.g., *might, could, may*), lexical verbs carrying epistemic value (e.g., *seem, think, believe, suggest, appear*), epistemic adjectives and nouns (e.g., *likely*, *possibility*), if-clauses and set expressions narrowing the universality of a claim (e.g., *to our knowledge*; *based on the model proposed here/based on this study*; Hyland 1998: 3-4; 45-46). In scientific writing, hedges are commonly used when the author hypothesizes interpretations of data, describes the results of a study, or states its limitations. Here, hedge words address what Skelton (1997) defines evidential truth (derived from the experimental study itself and the ensuing statistical judgment) and interpreted truth (the author's interpretation of findings), i.e., what the findings of a study are and what they mean (Hyland 1998: 56).

Hedges have been widely studied across different academic genres, and their use has been investigated among native and non-native speakers of English (henceforth, NSs and NNSs; see Hyland and Milton 1997; Hinkel 2005; Aull and Lancaster 2014; Vandenhoek 2018). In this context, hedges are often opposed to boosters, another metadiscursive strategy employed in academic writing to voice a stronger level of commitment. Boosters are also expressions of stance and engagement with the audience (Hyland 2005: 179), but they "[increase] epistemic commitment" and express certainty (e.g., adverbs *clearly* and *definitely*; Aull and Lancaster 2014: 160) – a function that is nevertheless common in academic writing. While hedging weakens claims, boosting strengthens them, and where hedges "open dialogic space, [by] leaving space for alternatives or disagreement, […] boosters such as *certainly* or *definitely* close dialogic space: They show full commitment or certainty and leave little room for disagreement" (Aull 2019: 270).

In what follows, we will explore how some hedging devices are used by advanced EFL learners in a specific, under-explored academic writing genre, i.e., thesis abstracts. The writing of L2 thesis abstracts is a required academic skill for language specialist students at undergraduate and graduate levels in Italian universities (majoring, for example, in foreign languages and linguistics, linguistic and cross-cultural mediation, or translation and interpreting studies). However, abstracts may be somewhat neglected as a genre in traditional academic writing courses, where students' attention is drawn to other, more common text types (e.g., essays and reports). To cover this knowledge gap, this study investigates the use of hedging in abstract writing within a corpus of texts produced by C1/C2-level Italian EFL university students, with a special focus on a set of hedge words (see Section 3.2).

2. Hedging in L2 academic writing

Hedging is a complex area to master in L2 academic writing. First of all, using hedges appropriately (or at all) is part of learners' academic writing skills, but especially pertains to pragmatic competence, which is typically late acquired (Fields and Matsuda 2018; Taguchi 2018). Secondly, hedge words, like many pragmatic markers, are usually polyfunctional expressions with no univocal form-meaning correspondence – and different hedges can serve the same function (Hyland 1998; Fraser 2010). This lack of one-to-one correspondence makes it difficult for learners to extract and generalize usage patterns. As a higher-level rhetorical device, hedging is not always foregrounded or is only partially covered in academic writing textbooks and courses (see, for example, Holmes 1988; Hyland 1994, 2008).

In the literature, EFL learners from different L1 backgrounds and disciplines have been seen to generally use fewer hedges than NSs in their academic written production, which often ends up sounding more assertive (Hyland and Milton 1997; Hinkel 2005; Siew Mei 2007; Vandenhoek 2018). The trend has been recorded with both upper-intermediate and advanced learners, although the latter progressively select more hedges over boosters as L2 proficiency increases (Siew Mei 2007; Aull and Lancaster 2014). L2 learners also generally tend to employ a narrower range of hedges as compared to NSs in essays and theses (Hyland and Milton 1997; Hyland 1998; Abdollahzadeh 2011). Most privileged forms often include modal

verbs, which are typically given more emphasis over other hedging devices in academic writing textbooks (Hyland 1994, 1998; Demir 2018).

Although similar results have been found across different student groups, it is to be noted that different L1s and cultures adhere to diverse rhetorical models and practices, which may affect the directness and assertiveness of claims in academic discourse (Mauranen 1993; Hinkel 2005; and more generally on metadiscursive functions Ädel 2008). In parallel, different disciplines make different use of hedges and stance markers in general. Hedging appears to be generally more pervasive in soft vs. hard sciences (Hyland 2005), although differences have been observed when specific subgenres are concerned (an example is Hyland and Zou 2021's work on three-minute theses, where more hedges happened to be used by students in the hard sciences; see also Aull 2019). Teaching-oriented studies have shown that hedging, like pragmatic competence in general, appears to benefit from explicit instruction in L2 academic writing classes (Wishnoff 2000; Algi 2012). Camiciottoli (2003) carried out an exploratory study on metadiscourse and ESP reading comprehension of L1-Italian college students and highlighted the need for targeted instruction in that regard. These studies thus suggest the need for both textbooks and instructors to focus more extensively on the expression of stance in academic discourse.

Within research on hedging in EFL academic writing (see, for example, Feng 2019; Firoozjahantigh et al. 2021; Zhao and Liu 2021), there is still little on texts produced by Italian natives. Neff et al. (2003) examine the use of evidential hedges in the writings of NS and NNS with five different L1s (Italian was among those). The comparison between L1 and L2 shows more limited use of lexical modality expressions in NNS writings, but the research tackles the specificity of Italian learners only partially. Dheskali (2020) analyzes four corpora consisting of L1 (Italian and Albanian) and L2 (Italian-English and Albanian-English) academic writings of college students to see what choices students make in the usage of hedges and boosters. Overall, Italians appear to use more boosters than Albanians, both in their L1 and L2, although slightly more in the former, similarly to Hyland's (2003) findings which showed that boosters tended to have prevalence in the L1. This study attempts to contribute to the still little research on hedges and epistemic markers produced by Italian learners in academic writing and it does so by investigating a very specific genre: abstracts.

3. The study

3.1. Aims and research questions

The current study focuses on the use of hedges by L1-Italian advanced learners of English within thesis abstracts, an academic genre that has not so far been extensively addressed in the literature. In general, abstracts are short summaries/accounts of a longer text, typically a research article, a thesis, or a dissertation. Thesis abstracts are common "in institutions where theses and dissertations are written in the local language [and] the accompanying English abstracts can be used to showcase departmental research to the wider world" (Swales and Feak 2009: 4) or serve as an additional aspect of evaluation when specialist students are concerned. This is a common situation in Italian universities, especially with students majoring or graduating in foreign languages, linguistics and translation. Although different studies have addressed hedging in academic theses and dissertations (e.g., Abdollahzadeh 2011), L2 thesis abstracts still remain quite an unexplored area, probably given their high cultural specificity and their limited geographical diffusion (e.g., in some European countries; Swales and Feak 2009).

Research carried out on conference abstracts typifies abstracts as progressing along five discourse moves, consisting of introduction, purpose, method, results, and conclusion (Egbert and Plonsky 2015; Ioratim-Uba 2020). Thesis abstracts follow the same structure, aligning with the typical IMRD format (Introduction, Method, Results and Discussion) on which research articles are based. Given the emphasis on the novelty and significance of a study, the presentation and discussion of its results and its framing within existing research, hedging is also a key component of abstracts (Ioratim-Uba 2020) and is an important feature to be mastered by L2 students even when this academic genre is concerned.

Moving from these premises, the study addresses the following research questions:

i. To what extent are hedging devices used in thesis abstracts by L1-Italian EFL advanced university undergraduates specializing in cross-linguistic mediation and foreign languages? How pervasive is the use of hedges as opposed to boosters?
ii. How do our trends compare with texts produced by L1-English senior undergraduates from similar disciplines?

The use of boosters was compared to the occurrence of hedges in the corpus since the former sometimes prevail in students' writing (Hinkel 2005; Aull and Lancaster 2014), although some research shares different results (see for example Dheskali 2020).

In order to provide answers to the research questions, the study is divided into two phases, which will be described in what follows.

3.2 Method and corpora

The study is corpus-based and draws on two different small corpora. The first phase of the research aims at exploring hedging strategies in thesis abstracts written by advanced EFL university undergraduates. A corpus was created including 217 English-language Bachelor's thesis abstracts produced by third-year university students about to graduate in Linguistic and cross-cultural mediation, and consisting of 287,048 words (henceforth, the Thesis Abstract Corpus or TAC). The undergraduate students had an advanced level of competence in L2 English (approximating C1/C2 CEFR level) and had attended English academic writing classes during their university career. Being this a group of specialist students, a foreign language abstract is required as an integral part of their theses – when such theses are written in the students' native language, i.e. Italian. The theses covered a wide range of topics in the Humanities, specifically in the subfields of Translation Studies, Linguistics, Cross-cultural Mediation, Foreign Language Acquisition, Literature, Art history, History and Geography.

A set of hedges was singled out and investigated in the corpus. Hedge words were selected based on frequency of occurrence in written production as emerged from previous research (Hyland 1998; Aull and Lancaster 2014; Aull 2019) and distributed over different word classes. Their main pragmatic function is to mitigate the assertive force of statements and, to a lower extent, to avoid overgeneralization (cf. Aull et al. 2017). The target hedges included:

- modal verbs of tentative epistemic possibility *might/may/could*;
- four lexical verbs commonly used as hedges and divided into two categories following Hyland's terminology: speculative judgmental verbs *suggest* and *indicate* and sensory evidential verbs *seem* and *appear* (Hyland 1998: 126);

– epistemic stance adverbs *possibly, generally, in general*, *slightly*, and *perhaps* (cf. Aull and Lancaster 2014; Aull 2019);
– epistemic adjectives *possible* and *likely*.

By way of comparison, a set of boosters was also investigated (in line with previous studies, e.g. Aull and Lancaster 2014) and included:

– modal verbs *can* and *must*; epistemic *will* was also added;
– adverbs *certainly, absolutely, definitely, clearly* and *undoubtedly*;
– the verb *demonstrate*.

In academic writing, *can* expresses a stronger, logical possibility (e.g. *These observations can be explained biochemically*; Biber et al. 1999: 492), and was included in the analysis given its frequency in NNSs' written production (Hyland 1994, 1998; Demir 2018). Epistemic *must* indicates inferential certainty (Coates 1983; Hyland 1998: 106). Epistemic *will* was additionally included in the analysis despite not being classified as a booster, as it expresses a confident epistemic prediction (Hyland 1998).

A second phase of the research involved a comparison of the use of the target hedges and boosters with NSs' academic written production. The comparison was drawn by searching a subsection of the MICUSP Corpus (Römer and Swales 2010), specifically filtering the query by student level, NS status, discipline and written genre (searchable database at https://elicorpora.info/main). In particular, the subsample consisted of written texts produced by senior undergraduates whose L1 was English; written genres included argumentative essays, critiques/evaluations, proposals, reports and response papers in the fields of Education, English and Linguistics. Since abstracts are not included in MICUSP, roughly comparable genres in terms of purpose or structure were selected. However, given the lack of fully comparable texts, the outcomes of the comparison will be taken with due caution. The final subsample from MICUSP consisted of 214,900 words for a total of 105 papers.

The two corpora were searched individually through the concordancing tool on Sketch Engine (Kilgarriff et al. 2014; http://www.sketchengine.eu). Our findings for different metadiscourse categories and items are illustrated in the following sections.

3.3 Analysis and results

3.3.1 Data from the *Thesis Abstract Corpus*

The first phase of the study explored the use of hedges in the Thesis Abstract Corpus. Table 1 shows the frequency of occurrence of different hedges in the corpus, differentiated by category. Raw frequency values are followed by normalized frequency per million words (p.m.w.).

Hedge category	Specific hedge	Raw frequency	Normalized frequency (per million words)
Modal verb (tentative possibility)	*may*	104	362.31
Modal verb (tentative possibility)	*might*	53	184.64
Modal verb (tentative possibility)	*could*	222	773.39
Epistemic stance adverb	*possibly*	6	20.9
Epistemic stance adverb	*generally*	10	34.84
Epistemic stance adverb	*in general*	11	38.32
Epistemic stance adverb	*slightly*	3	10.45
Epistemic stance adverb	*perhaps*	10	34.84
Epistemic adjective	*likely*	11	38.32
Epistemic adjective	*possible*	96	334.44
Epistemic lexical verb, judgmental, speculative	*suggest*	38	132.38
Epistemic lexical verb, judgmental, speculative	*indicate*	17	59.22
Epistemic lexical verb, evidential, sensory	*seem*	72	250.83
Epistemic lexical verb, evidential, sensory	*appear*	25	87.09

Table 1: Target hedges in TAC

Modal verbs are the most common hedging strategies in the corpus, confirming findings from previous research (Hyland 1994, 1998). *Could* prevails (773 occurrences p.m.w.), followed by *may* (362 hits p.m.w.)

and, to a lower extent, *might* (53 tokens and 185 occurrences p.m.w.) as expressions of tentative possibility (Examples 1, 2, 3 below). Other extremely popular hedges used by students in the corpus are the adjective *possible*, with 96 epistemic uses in TAC (334 p.m.w.) and the epistemic verb *seem* (251 occurrences p.m.w., Example 4). In its epistemic meaning, *possible* mostly occurs as an attributive adjective (49 occurrences, 170.7 p.m.w., e.g. *possible solutions*, *possible method*) and in the *n*-gram *it is possible to* (39 occurrences, 135.87 p.m.w., Example 5):

(1) This *could* be one of the motivations of the incredible sudden success of the artist, the same success that led him to a premature death.
(2) Hard of hearing people interact through spoken language and they *may* benefit from assistive devices like cochlear implants.
(3) there *might* be a problem of incompatibility between the model of digital device and learning software, browser or application.
(4) it is a condition in which all the women in her stories are forced to live, from childhood to death, there never *seems* to be an escape
(5) In fact, *it is possible to* declare that translation of vulgarity is mainly influenced by functional and semantic correspondance [sic] between source and target language.

Epistemic stance adverbs do not frequently appear in the corpus. When they do, they occasionally combine with other hedges, as is the case for *possibly* (2 occurrences out of 7 in combination with modals, especially *can* and *would*, Example 6). The trend is consistent with findings in Hyland (1998: 150), where the adverb *possibly* was a common trigger for hedge clusters.

(6) Aforementioned study enabled to obtain the inescapable framework by means of whom it was feasible to devise an educational system that *would possibly* be as appropriate as possible for such learners/the subjects treated. [sic]

Other hedge clusters are built around epistemic lexical verbs, as illustrated in Examples 7, 8 and 9. In particular, the clusters *may seem* and *could appear* occur 3 times each in the corpus (10.45 hits p.m.w.).

(7) Sometimes it *may seem* impossible and people may feel small in front of a strong system which oppress and homologate everybody

(8) it should be stressed that the integration and interdependence between the two approaches cannot be ignored, despite the fact that they *may appear* antithetical at first glance.
(9) intelligence *might indicate* a specific species, which homo sapiens belongs to

The modal verbs *can*, *will*, and *must* are additionally included in the analysis, and express logical possibility or greater certainty and confidence in a prediction or inference (Biber et al. 1999). Table 2 shows their frequency in the corpus and p.m.w.:

Modal verb	Raw frequency	Normalized frequency (per million words)
can	590	2,055.41
must	6	20.9
will	205	714.17

Table 2: Modals can, must *and* will *in TAC*

Can is widely used in the Thesis Abstract Corpus to express logical possibility (Example 10), especially in the passive voice, while *must* does not frequently appear in its epistemic meaning. *Will* is often used to express commitment and introduce further phases in the research/abstract, but also shows 205 epistemic occurrences where it voices a confident prediction (714 tokens p.m.w., Example 11):

(10) Hence, the translation activity *can* be defined as a matter of choice and a negotiation of meanings
(11) automatic translation *will* always continue to have weaknesses compared to the classical translation process.

For the sake of comparison, a set of additional boosters was also included in the study. Table 3 illustrates absolute and normalized frequency values.

Category	Specific booster	Raw frequency	Normalized frequency (per million words)
Adverb	*certainly*	17	59.22
Adverb	*absolutely*	2	6.97
Adverb	*definitely*	20	55.74
Adverb	*clearly*	29	101.03
Adverb	*undoubtedly*	18	62.71
Lexical verb	*demonstrate*	74	257.8

Table 3: Boosters in TAC

Adverbial boosters expressing higher certainty show generally higher frequency than adverbial hedges expressing tentativeness and indicate a preference in the corpus for more assertive epistemic adverbs. Some examples are shown in 12, 13 and 14.

(12) The role of English, in the 21st century, is *clearly* undisputed
(13) the key-points of his method – which he called "Talent Education" – *definitely* show his will to create the right environment for learning music
(14) The identification of the speaker on the screen is *certainly* problematic for deaf people

Demonstrate also prevails over the single frequencies of the epistemic lexical verbs queried within the corpus – *seem, appear, suggest* and *indicate* (Example 15):

(15) The results gathered *demonstrate* that the mechanisms which operate in the novel are complex and that the novel cannot be interpreted uniquely

After showing data from MICUSP, we suggest a comparison between hedges and boosters within the two corpora.

3.3.2 Data from MICUSP

A parallel search in the MICUSP corpus sections shows the following results (Table 4):

Hedge category	Specific hedge	Raw frequency	Normalized frequency (per million words)
Modal verb (tentative possibility)	*may*	240	964.66
Modal verb (tentative possibility)	*might*	100	401.94
Modal verb (tentative possibility)	*could*	253	1,016.91
Epistemic stance adverb	*possibly*	13	52.25
Epistemic stance adverb	*generally*	28	112.54
Epistemic stance adverb	*in general*	13	52.25
Epistemic stance adverb	*slightly*	10	40.19
Epistemic stance adverb	*perhaps*	76	305.48
Epistemic adjective	*likely*	26	104.5
Epistemic adjective	*possible*	49	196.95
Epistemic lexical verb, judgemental, speculative	*suggest*	83	333.61
Epistemic lexical verb, judgemental, speculative	*indicate*	24	96.47
Epistemic lexical verb, evidential, sensory	*seem*	276	1,109.36
Epistemic lexical verb, evidential, sensory	*appear*	67	269.3

Table 4: Target hedges in MICUSP

Even in the MICUSP corpus sections, modal verbs are a common hedging device, following the same progressive distribution as in TAC, but at a higher frequency. Examples 16, 17 and 18 show some uses in context of the three verbs:

(16) *Could* their plots be translated without contradicting the narrative's style?

(17) A problem *may* be that the students do not think the participation grade will have a strong effect on their grade.
(18) However, in the spirit of conducting analysis rather than writing review, a better starting point *might* be to question the possible motivations underlying directorial choices

Epistemic stance adverbs and adjectives are also common, especially *perhaps, generally*, *possible* and *likely* (Examples 19 to 22).

(19) As the poem continues, it becomes apparent that this case is *perhaps* indicative of love in general in some respects.
(20) This *generally* ensures that the students are able to meet a good amount of the state standards for social studies in middle schools.
(21) It is *possible* that one of the teachers still currently on staff has kept up with that teacher.
(22) Jim's negative attitude is *likely* to cause further frustration for both Carlos and himself

Seem is the most frequent lexical verb carrying epistemic value (276 hits in the corpus, 1,109 p.m.w.), followed by *appear* (67 occurrences, 269 p.m.w.) and, with way fewer occurrences, *indicate* (24 tokens, 96 p.m.w.). Examples 23, 24 and 25 illustrate some concordances for the three verbs:

(23) He didn't *seem* to believe in himself at the end either.
(24) On that account, it *appears* as though Bloom intends to keep the affair innocuous;
(25) His behavior *indicates* a sort of war of trauma victims

Similarly to what was observed in the Thesis Abstract Corpus, hedge clusters appear around epistemic stance adverbs and lexical verbs. Most common n-grams include *would seem* (14 tokens, 56.27 p.m.w.), *may seem* (3 tokens, 12.06 p.m.w.), *could possibly* (3 tokens, 12.06 p.m.w.).

For better comparison, Figure 1 shows the frequency of hedges across the two corpora normalized per million words:

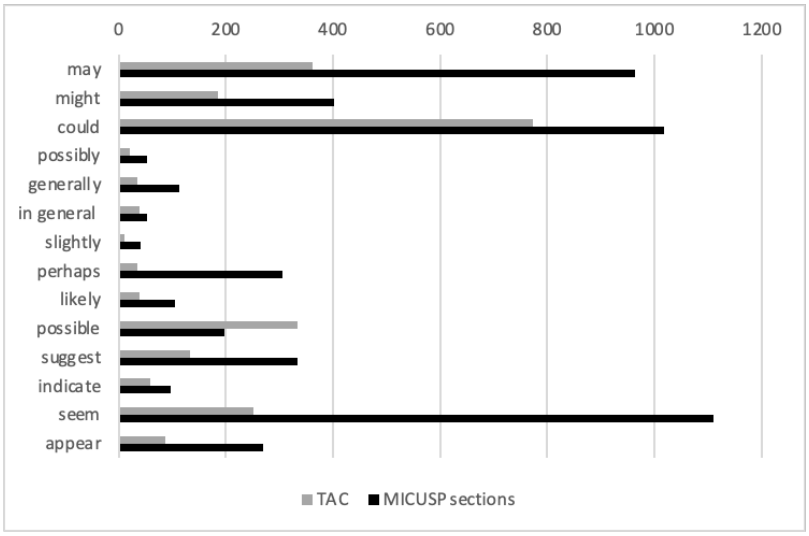

Figure 1. Frequency of hedges in the Abstract and MICUSP corpora (p.m.w.)

As is evident from Figure 1, hedges are more common in the MICUSP Corpus sections. The trend concerns all categories of hedges, and is especially striking for modal verbs, lexical verbs *seem* and *suggest*, epistemic stance adverbs and the epistemic adjective *likely*. The single exception is the adjective *possible* used to express tentativeness, which seems to be overused by Italian EFL learners – a trend that may be partly related to its Italian cognate *possibile* (Section 4).

Chi square statistics show a significant difference between the two corpora for all categories, i.e. in the use of epistemic modal verbs ($\chi2=485.98$, $p<.00001$), epistemic stance adverbs ($\chi2=52.86$, $p<0.00001$), epistemic stance adjectives ($\chi2=60.12$, $p<.00001$) and lexical verbs ($\chi2=43.92$, $p< 0.00001$).

The modal verbs *can*, *will*, and *must* are used as follows in the MICUSP Corpus (Table 5):

Modal verb	Raw frequency	Normalized frequency (per million words)
can	393	1,579.63
must	18	72.35
will	312	1,254.06

Table 5: Modals can, must, *and* will *in MICUSP*

Modals *can* and *will* are the most frequent (examples 26 and 27), while *must* only occurs 18 times (72 occurrences p.m.w.), paralleling data from TAC. The similar distribution of these three modal verbs, despite their different frequency, is better illustrated in Figure 2. While *must* and *will* prevail in MICUSP, *can* is more popular in TAC, reflecting the tendency for NNSs to overuse this modal when expressing possibility. In terms of frequency, the two corpora differ significantly in the use of the three modal verbs in question ($\chi 2=235.94$, $p< 0.00001$).

(26) this sentence *can* also be read as a lens through which to interpret the narrative style.

(27) This activity not only *will* help develop her students' literary skills, it also teaches her students how to be responsible.

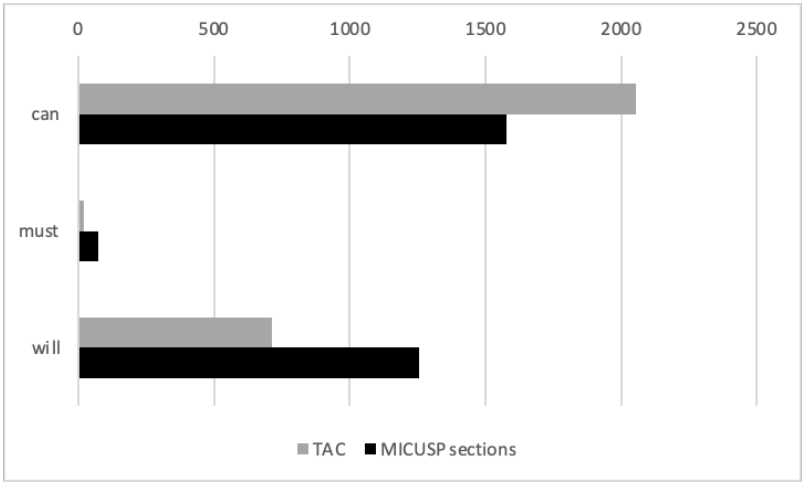

Figure 2. *Can*, *must*, and *will* in the TAC and MICUSP corpora (p.m.w.)

When looking at boosters, the following trends emerge in the MICUSP Corpus (Table 6):

Hedge category	Specific hedge	Raw frequency	Normalized frequency (per million words)
Adverb	*certainly*	38	152.74
Adverb	*absolutely*	4	16.08
Adverb	*definitely*	11	44.21
Adverb	*clearly*	36	144.7
Adverb	*undoubtedly*	10	40.19
Lexical verb	*demonstrate*	58	233.13

Table 6: Boosters in MICUSP

Boosters generally occur in the corpus to a lower extent than hedges. Among the categories analyzed, the most common ones are the verb *demonstrate* (58 hits, 233 occurrences p.m.w.) and the adverbs *certainly* (38 occurrences, 153 p.m.w.) and *clearly* (36 tokens, 145 p.m.w.). Although hedging is a more pervasive function, boosting is also a common strategy in academic writing among L1-English senior undergraduates (examples 28, 29 and 30).

(28) Bhabha's concept of mimicry *demonstrates* one of the inherent contradictions in colonialism

(29) Hosea's metaphor *certainly* reestablishes the idea of the covenant as one of love

(30) Language policy and education clearly *shows* that for now, this so-called melting-pot society is actually a coercive assimilation

A direct comparison between the two corpora shows the following trends (Figure 3):

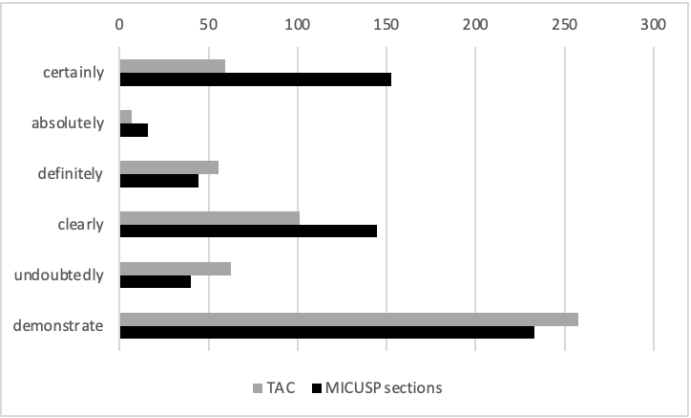

Figure 3. Frequency of boosters in the TAC and MICUSP corpora (p.m.w.)

Adverbs *certainly, clearly* and *absolutely* prevail in MICUSP. Conversely, adverbs *undoubtedly* and *definitely* and the lexical verb *demonstrate* are slightly more pervasive in the Thesis Abstract Corpus.

Overall, there is a significant difference in the use of adverbial boosters ($\chi 2=42.45$, $p< 0.00001$), while the different frequency of the lexical verb *demonstrate* in the two corpora is not statistically significant ($\chi 2=1.0221$, $p=.31$).

4. Discussion

The quantitative analysis of the Thesis Abstract Corpus shows that hedging strategies are used by L1-Italian, advanced, specialist EFL students in their thesis abstracts. The most common devices that appear in the corpus to express tentativeness and lower certainty are modal verbs, with a predominance of *could*. *May* follows, but shows about half of the occurrences of *could*. *Might* is found on occasion, but is not as popular as its counterparts. Within our set of hedges, verbs are the privileged means through which hedging is achieved, since modals are followed in frequency by epistemic lexical verbs, among which the most commonly used is evidential sensory *seem*, followed by judgmental speculative *suggest*. Most of the epistemic verbs we considered (*suggest, indicate, appear*) are

of Latinate origin, and this adds up to the reasons why they may be easily adopted by Italian students. The adjective *possible* is also frequently found in the Thesis Abstract Corpus carrying epistemic value, and it appears in two preferred syntactic constructions, i.e. as a noun premodifier and as part of the *n*-gram *it is possible to*. The Italian cognate adjective *possibile* can be used in similar constructs – as it can take premodifying position when it is used epistemically and commonly activates the impersonal construction *è possibile + che / è possibile* + infinitive ('it is possible that'/'it is possible to'; Pietrandrea 2004). Although adverbs can carry epistemic functions in Italian, they do not occur frequently in the corpus, and are rarely adopted by the students as hedging devices.

While advanced EFL learners do apply hedging when writing their thesis abstracts, and mitigate many of the assertions they make, they more often tend to express a higher degree of certainty in their statements or show greater confidence in their inferences. This claim aligns with findings from other studies on hedging and metadiscursive strategies (e.g. Hyland and Milton 1997; Abdollahzadeh 2011) and is coherent with the view of hedging varying as a function of L2 proficiency (Siew Mei 2007; Aull and Lancaster 2014).

When considering other modal verbs associated with higher certainty, *can* in its meaning of logical possibility is by far the most pervasive; many predictions are also formulated by using *will*. These modal verbs are core ones and are generally introduced early in EFL syllabi; in addition to that, they fulfill basic pragmatic functions – and their meanings may be more immediately associated with possibility with respect to other modals farther away on the *irrealis* continuum.[1] The tendency to higher assertiveness is further supported by the frequent use of the booster *demonstrate*. Like hedging ones, boosting adverbs are not very common in the corpus, but their occurrences still generally outnumber the former.

The comparison with NS data additionally confirms the trend. While hedges prevail in the MICUSP Corpus sections, many boosters occur more frequently in the Thesis Abstract Corpus. More in detail, all hedge types are more pervasive in the NS corpus and the difference is statistically significant for all categories – with hedges occurring in some cases twice or

[1] In many accounts, *can* is not always associated with an epistemic value (e.g. Hyland 1998), and even its speculative uses are associated with a subtype of dynamic modality (e.g. *A spider can be dangerous*, Portner 2009: 135–36). Strictly epistemic uses of *can* would only apply to its negative or interrogative form (Hyland 1998: 109).

three times as much in MICUSP. It is interesting to observe, however, that a similar data distribution is followed in the two corpora when looking at the use of modal and lexical verbs: in both TAC and MICUSP, the gradient of use is the same – *could-may-might* and *seem-suggest-appear-indicate*. Advanced L1-Italian EFL learners use fewer hedges than their L1-English peers, but distribute them proportionately (at least hedging verbs). A greater disparity in usage patterns is observed for adverbial and adjectival hedges, which are much more pervasive in NSs' writing. One notable exception is *possible*, whose frequency in TAC almost doubles that in MICUSP – and may be attributable to potential transfer from the L1.

While (greater) epistemic certainty is also widely pervasive in the MICUSP Corpus, and modals *will* and *must* prevail here, *can* is more popular in TAC. Once again, the gradient of frequency is the same (epistemic *can-will-must*), but the occurrences of *can* as a marker of logical possibility more than double those of *will* in NNSs' productions. Finally, typical boosters show a more diversified behavior: adverbs *certainly, absolutely,* and *clearly* are significantly more frequent in MICUSP; verb *demonstrate* and adverbs *definitely* and *undoubtedly* are more popular among EFL learners – but the difference in use is not statistically significant for *demonstrate*. It is thus fair to say that abstracts written by NNSs still privilege epistemic certainty over hedging, and may tend to sound more assertive, but assertiveness is also typical of NSs' academic writing.

While the comparison with NSs' behavior may be useful to analyze and contextualize our data, the results have to be interpreted with due caution, given that the written genres in the two corpora are not fully comparable, and this could have introduced further elements of variation. As emerged in some studies, genre may be a relevant factor in determining the use and distribution of hedges (Aull 2019; Hyland and Zou 2021).

5. Conclusion

Hedging is a key aspect of academic writing, as it contributes to the reliability of a study and opens a dialogic relation with readers and the scientific community at large. Hedging devices, as well as the degree to which hedge words are used, may vary cross-culturally and often represent a challenge for L2 learners. In exploring EFL learners' use of hedges in thesis abstract writing, our study generally confirms this. Different categories of

hedge words are found in the Thesis Abstract Corpus, but are not generally as frequent as boosters or modal verbs expressing higher certainty in claims, inferences and predictions. Hedges are also less common and less diversified than those used by NSs as recorded in the MICUSP sections, thus showing parallel results to similar studies on hedging and the use of metadiscursive strategies by NNSs and L2 learners (Hyland and Milton 1997; Hyland 1998; Abdollahzadeh 2011).

Preferred hedges in learners' written productions include modals *may* and *could*, some non-factive verbs expressing evidentiality and speculation, and the epistemic adjective *possible*. Although, with the exception of *possible*, these hedges are still underrepresented compared to boosters or NSs' writing, their privileged status may be linked to their prototypicality as hedging devices and their similar use in the learners' L1. The hedging function is more immediate with tentative epistemic modals, and is the main meaning associated with modals *may* and *could*. The same could be true for verbs like *seem* and *suggest*, whose evidential and speculative value is semantically prominent. In addition to being similarly prominent, the epistemic uses of *possible* in premodifying position or in the pattern *it is possible to* also parallel equivalent structures in Italian, favoring learners' choice of this hedging device. The two factors (prototypicality and similarity to the L1) could have acted in combination in affecting the frequency of the adjectival hedge in TAC – given its pervasiveness in the academic production of writers from different L1s as well (Demir 2018).

In the Thesis Abstract Corpus, the modal *can* is also overused to express logical possibility and introduce a claim. The trend may be linked to the primacy of *can* as a modal verb – and aligns with similar results from previous studies (Abdollahzadeh 2011; Demir 2018).

The diversified use of both hedges and boosters in the two corpora points at a lower mastery of metadiscursive strategies in general by EFL learners in the Thesis Abstract Corpus, even though they are specialist students. Advanced learners did use hedging devices when writing thesis abstracts in English, but still showed high assertiveness when making and backing claims, and lower variability when selecting hedges and boosters alike. This may not always grant effective communication and may result in their writing sounding less reliable and too imposing on the audience. Metadiscursive devices thus represent a challenge even for learners at an advanced level of competence in L2 English (cf. Siew Mei 2007; Aull and Lancaster 2014), and call for more explicit instruction in academic

writing courses to help students navigate through the polyfunctionality and pragmatic complexity of these expressions and acquire full awareness of such complexity. Work on contrastive rhetoric and cross-cultural differences in hedging may be also beneficial in raising learners' awareness of the phenomenon and of its differentiated use across different rhetorical styles and conventions (Samaie et al. 2014; Chek Kim and Miin-Hwa Lim 2015), while at the same time defining varying stance expectations across disciplines and genres (Aull 2019).

The study is not deprived of limitations, as it draws on two small-sized corpora and does not grant full comparability of the results, especially in terms of academic genre. Further research might rely on larger writing samples and involve multimodal analyses of different metadiscursive strategies, in order to better characterize thesis abstracts as a discourse type of their own.

Acknowledgments

The article is the result of joint work by the authors. More in detail, Elisa Ghia wrote sections 3 and 5, Emilia Petrocelli wrote sections 1 and 2 and Sergio Pizziconi wrote section 4.

References

Abdollahzadeh, E. (2011). Hedging in postgraduate student theses: A cross-cultural corpus study, *International Conference on Languages, Literature and Linguistics*, 26, 581–586.

Ädel, A. (2008). Metadiscourse across three varieties of English: American, British, and advanced learner English. In: U. Connor, E. nagelhout and W. Rozycki (eds), *Contrastive Rhetoric. Reaching to intercultural rhetoric*. Amsterdam: Benjamins, 45-62. https://doi.org/10.1075/pbns.169.06ade

Algi, S. (2012). *Hedges and boosters in L1 and L2 argumentative paragraphs: Implications for teaching L2 academic writing* (Unpublished Master thesis). Middle East Technical University.

Aull, L. L. (2019). Linguistic Markers of Stance and Genre in Upper-Level Student Writing, *Written Communication*, 36(2), 267–295. https://doi.org/10.1177/0741088318819472

Aull, L. L. and Z. Lancaster (2014). Linguistic markers of stance in early and advanced academic writing: A corpus-based comparison. *Written communication*, 31(2), 151-183. https://doi.org/10.1177/0741088314527055

Aull, L. L. and D. Bandarage, M. R. Miller (2017). Generality in student and expert epistemic stance: A corpus analysis of first-year, upper-level, and published academic writing. *Journal of English for Academic Purposes*, 26, 29-41. https://doi.org/10.1016/j.jeap.2017.01.005

Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5, 97-116. https://doi.org/10.1016/j.jeap.2006.05.001

Biber, D. et al. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.

Boye, K. and P. Harder (2007). Complement-taking predicates: Usage and linguistic structure, *Studies in Language*, 31(3), 569–606. https://doi.org/10.1075/sl.31.3.03boy

Camiciottoli, B. C. (2003). Metadiscourse and ESP reading comprehension: An exploratory study. *Reading in a Foreign Language*, *15*(1), 28–44.

Coates, J. (1983). *The semantics of the modal auxiliaries*. Beckhenham: Croom Helm.

Demir, C. (2018). Hedging and academic writing: an analysis of lexical hedges, *Journal of Language and Linguistic Studies*, 14(4), 74–92.

Dheskali, V. (2020). *A Corpus-based Comparison of Albanian and Italian Student Writing in L1 and English as L2: Hedges and Boosters as Modalization by Degree*. PhD Thesis. University of Chemnitz.

Feng, J. (2019). A contrastive analysis of interpersonal metadiscourse in MA thesis literature review sections: The case of L1 Chinese students. In: M. Lehmann et al. (eds.), *UPRT 2017: Empirical Studies in English Applied Linguistics*. Lingua Franca Csoport, 291–309.

Fields, G. and P. K. Matsuda (2018). Advanced Rhetoric and Socially Situated Writing. In P. A. Malovrh & A. G. Benati (eds), *The Handbook of Advanced Proficiency in Second Language Acquisition*, Blackwell, Wiley, 527-546. https://doi.org/10.1002/9781119261650.ch27

Firoozjahantigh, M. and E. Fakhri Alamdari, A. Marzban (2021). Investigating the effect of process-based instruction of writing on the

IELTS writing task two performance of Iranian EFL learners: Focusing on hedging & boosting. *Cogent Education*, *8*(1), 1–15. https://doi.org/10.1080/2331186X.2021.1881202

Fraser, B. (2010). Hedging in political discourse. The Bush 2007 press conferences. In: U. Okulska and P. Cap (eds), *Perspectives in Politics and Discourse*, Amsterdam: John Benjamins, 201–214. https://doi.org/10.1075/dapsac.36.16fra

Garzone, G. (2006). *Perspectives on ESP and popularization*. Milan: CUEM.

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

Holmes, J. (1988). Doubt and certainty in ESL textbooks. *Applied linguistics*, 9(1), 21–44.

Hyland, K. (1998). *Hedging in scientific research articles*, Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.54

Hyland, K. (2003). *Second Language Writing*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511667251

Hyland, K. (2004). *Genre and second language writing*. Ann Arbor, MI: University of Michigan Press.

Hyland, K. (2005). Stance and engagement: a model of interaction in academic discourse, *Discourse Studies*, 7(2), 173–192. https://doi.org/10.1177/1461445605050365

Hyland, K. (2008). Make your academic writing assertive and certain. In: J. Reid (ed.), *Writing Myths*, Ann Arbor, MI: University of Michigan Press, 70–89.

Hyland, K. and J. Milton (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

Hyland, K. and H. J. Zou (2021). "I believe the findings are fascinating": Stance in three-minute theses, *Journal of English for Academic Purposes*, 50, 1-13. https://doi.org/10.1016/j.jeap.2021.100973

Ioratim-Uba, G. (2020). The impact of an instruction to improve L2 learners' academic conference abstract writing in multidisciplinary programmes. *Language Learning in Higher Education*, 10(1), 143–169. https://doi.org/10.1515/cercles-2020-2016

Kilgarriff, A. and V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý, V. Suchomel, (2014). The sketch engine: Ten years on. *Lexicography*, 1, 7–36. https://doi.org/10.1007/s40607-014-0009-9

Kim, L. C., and J. M. H. Lim (2015). Hedging in academic writing – a pedagogically-motivated qualitative study. *Procedia-Social and Behavioral Sciences*, 197, 600–607. https://doi.org/10.1016/j.sbspro.2015.07.200

Mauranen, A. (1993). Contrastive ESP rhetoric: Metatext in Finnish-English Economics texts. *English for Specific Purposes* 12, 3–22.

Neff, J.A. and E. Dafouz, H. Herrera, F. Martínez, J. P. Rica, M. Díez, R. Prieto, C. Sancho (2003). Contrasting learner corpora: The use of modal and reporting verbs in the expression of writer stance. In: S. Granger and S. Petch-Tyson (eds.). *Extending theScope of Corpus-based Research*: *New Applications, New Challenges*. Amsterdam: Rodopi. 211–230.

Pietrandrea, P. (2004), L'articolazione semantica del dominio epistemico dell'italiano. *Lingue e Linguaggio* 2, 171–206.

Portner, P. (2009), *Modality*. Oxford: Oxford University Press.

Römer, U., Swales, J. M., 2010, The Michigan Corpus of Upper-level Student Papers (MICUSP), *Journal of English for Academic Purposes* 9(3), 249.

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*. English Language Programs: United States Information Agency, 105–118.

Samaie M., Khosravian F., Boghayeri M., 2014, The Frequency and Types of Hedges in Research Article Introductions by Persian and English Native Authors, *Procedia – Social and Behavioral Sciences* 98, 1678–1685.

Siew Mei, W. (2007), The use of engagement resources in high- and low-rated undergraduate geography essays. *Journal of English for academic purposes*. 6 (3), 254–271.

Skelton J., (1997) The representation of truth in academic medical writing. *Applied Linguistics* 18(2), 121–140.

Swales, J. M. (1990). Genre Analysis: English in academic and research settings. Cambridge: Cambridge University Press.

Swales, John & Christine Feak. 2009. Abstracts and the writing of abstracts. Ann Arbor: University of Michigan.

Taguchi N. (2018), Advanced Second Language Pragmatic Competence in Malovrh & Benati (eds), *The Handbook of Advanced Proficiency in Second Language Acquisition*, Blackwell, Wiley, 505–526. https://doi.org/10.1002/9781119261650.ch26

Vandenhoek, T. (2018). Epistemic markers in NS and NNS academic writing. *Journal of Academic Writing*, 8(1), 72–91. http://dx.doi.org/10.18552/joaw.v8i1.359

Wishnoff, J. R. (2000). Hedging your bets: L2 learners' acquisition of pragmatic devices in academic writing and computed-mediated discourse. *Second Language Studies* 19(1), 119–148.

Zhao, J., & Liu, Y. (2021). A Developmental View of Authorial Voice Construction in Master's Thesis: A Case Study of Two Novice L2 Writers. *SAGE Open*, *11*(4), 1–15. https://doi.org/10.1177/21582440211054483

Received: 10 September 2022
Accepted for publication: 5 November 2022