

METADISOURSE IN L2 MASTER'S THESES: THE IMPACT OF ACADEMIC CULTURE AND EXPERTISE

*Olga Dontcheva-Navratilova, Tereza Guziurová, Renata Jančaříková
and Marie Lahodová Vališová*

Abstract

This study explores metadiscourse in English-medium Master's theses by L2 (Czech) graduates, aiming to explain how Czech students organise their texts, express their stance towards the content and engage with their readers. It seeks to contrast L2 learner academic discourse with L1 learner and expert academic discourse in order to identify differences along the culture and level of expertise dimensions. The corpus-based analysis employs Hyland's (2005) interpersonal framework of metadiscourse to identify the frequency, functions and realisations of interactive and interactional metadiscourse devices. The findings reveal that interactional metadiscourse is more prominent than interactive metadiscourse in all three corpora and there are significant differences in the realisation patterns and functions of specific metadiscourse markers. The results of the analysis suggest that self-mention, hedges and engagement markers vary along the expertise dimension as they are more heavily used in published research articles than in learner discourse. Cultural differences (i.e., those stemming primarily from different academic writing conventions) seem to affect the preferred degree of writer visibility, as well as preferences for specific metadiscourse markers. Variation in interactive metadiscourse seems to be influenced by text size, genre and communicative purpose. The findings allow for the drawing of several implications for L2 writing pedagogy.

Keywords

metadiscourse, Master's thesis, academic writing, genre, intercultural analysis

1 Introduction

Academic discourse involves a complex representation of socially contextualised knowledge and writer identity via language (cf. Duff, 2010, p. 175). When engaging in interaction with readers, academic writers attend to the interpersonal dimension of discourse signalled by "linguistic resources used to organize a discourse or the writer's stance towards either its content or the reader" (Hyland & Tse, 2004, p. 157). These interpersonal resources which allow the writer to guide readers through the discourse while seeking to persuade them to accept the writer's views and claims are subsumed under the concept of metadiscourse (Hyland, 2005). Views on what exactly falls into the scope of metadiscourse vary. Authors who consider reflexivity to be the defining feature of metadiscourse tend to restrict its scope to devices with text-organising functions and hold that metadiscourse does not contribute to the propositional content of the text; this approach is labelled 'non-integrative' or 'narrow' (e.g., Ädel,

2006; Mauranen, 1993). A different, 'broad' or 'integrative', approach is adopted by researchers who believe that metadiscourse is essentially interactional and distinct but inherently connected to the propositional aspects of discourse (Hyland & Tse, 2004; Hyland, 2004). This study adopts the integrative interpersonal metadiscourse model proposed by Hyland (2005), which comprises interactive devices centred on text organisation and interactional devices casting an authorial voice and engaging with the 'reader-in-the-text' (Thompson & Thetela, 1995).

Effective academic writing presupposes the skilful use of metadiscourse resources resonating with the shared beliefs, expectations and conventions of a specific academic community (Hyland & Jiang, 2018; Sancho-Guinda & Hyland, 2012). Learning to use metadiscourse effectively is not an easy task. This is particularly true for L2 university students, who have to cope with various language, rhetorical and psychological challenges stemming from their L2 writer status, cultural differences between academic writing norms and the pressure to perform well in high-stake examinations (Lee & Casal, 2014). The Master's thesis (MT), "the longest and most challenging piece of assessed writing" (Thompson, 2013, p. 284) that students compose at the end of their university studies, may be seen as indicative of the extent to which they have mastered the use of metadiscourse. Yet in comparison with the research article (RA) or the argumentative essay (AE), the MT has still received relatively little attention in research on academic writing (Lee & Casal, 2014; Thompson, 2013). This study seeks to fill in this research gap by exploring metadiscourse in L2 (Czech) students' MTs in the social sciences and humanities. By so doing, it aims at drawing pedagogical implications for academic writing courses to assist students in using the rhetorical potential of metadiscourse and thus enhance their academic writing competence.

Previous investigations into metadiscourse in English-medium RAs by L2 scholars (Lorés Sanz, 2011; Shaw, 2003) have found that the use of these rhetorical features is marked by intercultural variation concerning primarily the frequency of occurrence of individual devices, but also by the preference towards specific lexico-grammatical patterns. This could be interpreted as an indication of hybridisation in the English-medium discourse of L2 scholars, who seem to blend the discursive and rhetorical conventions of their original academic literacy with those of Anglophone academic discourse (Pérez-Llantada, 2013).

Resolving the tension between L1 and L2 academic norms is even more challenging for L2 students; however, intercultural variation in the use of metadiscourse in L2 learner academic writing has received considerably less attention and the existing studies focus mainly on the Asian context (e.g., Ho & Li, 2018; Hyland, 2004; Lee & Deakin, 2016). Therefore, this study endeavours to provide an insight into intercultural differences between the Anglophone and

L2 (Czech) academic writing conventions to raise the students' awareness of the existing differences and assist them in making informed rhetorical choices in their academic texts.

Several investigations exploring undergraduate AEs have found that the differences they exhibit in comparison to academic writing norms tend to be induced by L1 transfer, coping strategies, overgeneralisation, input bias and disagreement between instructions provided by style manuals and common practice in expert academic discourse (Crosthwaite et al., 2017; Hong & Cao, 2014; Qin & Ucelli, 2019). Studies comparing high-graded and lower-graded essays have revealed that successful essays mostly show metadiscourse patterns somewhat similar to expert discourse (Ho & Li, 2018; Lee & Deakin, 2016). The few studies exploring metadiscourse in MTs tend to adopt quantitative methods and focus primarily on interactional metadiscourse (Hyland, 2004; Lee & Cassal, 2014; Liu & Zhang, 2022; Qiu & Ma, 2019; Wu & Paltridge, 2021). Typically comparing MTs to doctoral theses or RAs, they show that L2 graduates underuse or overuse specific metadiscourse categories, and they report a developmental trend across levels of expertise. This points to the need to explore differences in the way writers of different levels of expertise use metadiscourse and tailor academic writing instructions to the needs of university students to provide scaffolding for the development of their writing skills.

Research into Czech English-medium learner discourse is restricted to a few studies focusing on specific metadiscourse markers, namely, sentence linkers (Povolná, 2012; Vogel, 2008), code glosses (Guziurová, 2022), attitude markers (Jančaříková, 2023), self-mentions (Dontcheva-Navratilova, 2023a) and endophoric markers (Lahodová Vališová, 2024). Obviously, the insights gathered from these studies are not sufficient to provide a comprehensive picture of the use of metadiscourse by Czech university students in comparison to L1 learner and expert academic discourse. Therefore, the aim of this study is to carry out a contrastive analysis between the use of the devices in L2 Czech learner discourse and L1 learner and expert discourse to see whether and to what extent they differ along the cultural and expertise dimensions.

The present study adopts a mixed-method approach to answer the following research questions:

- 1) Are there significant differences in the frequency of occurrence of metadiscourse devices in English-medium L2 (Czech) MTs, L1 AEs and L1 published RAs?
- 2) What are the differences and similarities in the realisations and functions of metadiscourse markers in L2 (Czech) learner discourse and L1 learner and expert discourse?

2 Data and method

2.1 Corpus

The study is based on a specialised learner corpus of English-medium MTs written by Czech university students (the MT corpus). The authors are L1-Czech postgraduate students majoring in English Language and Literature at Masaryk University in Brno, Czech Republic. The MT corpus consists of 48 theses in the domains in which students pursuing a Master’s degree in English Language and Literature typically write their theses – linguistics, literature studies, and education (16 texts per discipline). All the theses were defended between 2010-2018 and graded ‘A’ (‘Excellent’/‘Merit’) to match the quality of works in the reference L1 learner corpus. Prior to the analysis, the texts were processed to exclude citations, examples, tables, charts, and reference lists to ensure a focus on the students’ own discourse. As a result, the corpus used for analysis contains a total of 947,492 words. However, the analysis of citations within the evidentials category was carried out on the full length of the texts.

In order to explore the typical metadiscourse features used by Czech university students, two reference corpora were compiled following the principles of ‘tertium comparationis’, that is, creating corpora on the basis of relevant similarity constraints (Connor & Moreno, 2005). The two corpora include an L1 English learner corpus to investigate potential variation across linguacultural backgrounds, and a corpus of published RAs intended for comparison along the expertise dimension.

Due to the unavailability of an MT corpus in L1 English that would represent the target disciplines, and since it proved impossible to compile such a corpus because of access limitations, we decided to use a portion of the *British Academic Written English Corpus* (BAWE) for analysing variations between the English-medium learner discourse of L1 English writers and Czech graduates. In order to guarantee maximum comparability between the MT corpus and the learner reference corpus, a BAWE sub-corpus was created, consisting of 197 AEs authored by L1 English students, which, similarly to the theses, received grades of ‘Distinction’ or ‘Merit’. The essays belong to the disciplinary group of Arts and Humanities and encompass similar fields to those in the MT corpus, namely linguistics, English literature, and comparative American studies (as a substitute for education, which is not represented in BAWE). Although AEs and MTs differ in length, with essays averaging 2,500 words and theses 19,700 words, and partly in their communicative purposes, both genres share the common context of an examination setting. Therefore, we believe that this makes the BAWE “an ‘analogue’ corpus, that is, a corpus which is as near as possible in terms of genre

and discipline” (Flowerdew, 2015) to the MT corpus, and can thus be considered an acceptable reference corpus. Following the same data cleaning procedure, the BAWE corpus consists of 490,874 words.

The second reference corpus (RA) comprises 36 RAs written by L1 English authors (17 British, 17 American and 2 Australian) and is fully comparable with the MT corpus in terms of disciplines, being represented by linguistics, literature, and education (12 RAs per discipline). The articles were selected from well-established academic journals indexed in the Web of Science database (3 journals per discipline). All the texts were published between 2010 and 2018 and are single-authored. While RAs and MTs obviously differ in their communicative purposes, audiences and requirements to be met (Paltridge, 2002), they are both research-process genres with “significant areas of overlap in lexico-grammar and rhetorical functions” (Flowerdew, 2015, p. 60). As Swales (1990, p. 178) remarks, some chapters of theses or dissertations may later appear as RAs. Having been cleaned, the RA corpus contains a total of 242,439 words. Table 1 shows the composition and size of the corpora.

Corpora	Texts	Wordcount	Disciplines
MT	48	947,492	Linguistics, Literature, Education
BAWE	197	490,874	Linguistics, Literature, Comp. American Studies
RA	36	242,439	Linguistics, Literature, Education

Table 1: Composition of the MT, BAWE and RA corpora

2.2 Analytical framework and procedure

This investigation adopts Hyland’s (2005) interactional metadiscourse framework, which comprises two types of metadiscourse categories – interactive and interactional – differentiated according to the functions they fulfil. Interactive metadiscourse is associated with the Hallidayan textual metafunction; it helps the writer build the argumentation chain and navigate the reader through the text, thus enhancing discourse coherence and facilitating text comprehension. As an instantiation of the Hallidayan interpersonal metafunction, interactional metadiscourse projects the writer’s views and evaluative opinions into the text and engages in a dialogue with readers with a view to persuading them to accept the writer’s claims.

Interactive metadiscourse involves the following categories:

- transitions – indicate logical relations between main clauses and sentences (*then, however, thus*)
- frame markers – signal discourse organisation and argument development (*firstly, to summarise*)

- endophoric markers – indicate intratextual relations (*see Table 2, as noted in section 2*)
- evidentials – refer to sources of information outside the text (*X argues, according to Z*)
- code glosses – provide reformulations and examples to assist the reader in comprehending the text (*i.e., e.g., namely*).
- Interactional metadiscourse also comprises five categories:
- self-mentions – indicate authorial presence by exclusive personal pronouns (*I/we*) and possessives (*my/our*)
- attitude markers – express the writer's feelings and evaluative assessment (*valuable, significant, important*)
- hedges – express caution and reduce commitment to views and claims and invite alternative views (*typically, possibly, may*)
- boosters – enhance certainty and close the dialogic space for negotiation of views (*in fact, certainly, no doubt*)
- engagement markers – appeal to readers, who are presented as peers following the unfolding argument (*you/your, of course, consider*).

The first four interactional metadiscourse categories are associated with expressing authorial stance, which is conceived as an attitudinal dimension projecting the author's voice and positions into the text to create a credible authorial persona. In contrast, engagement is perceived as an alignment dimension enabling the writer to construct the 'reader-in-the-text' (Thompson & Thetela, 1995), suggest intended interpretations and signpost the argument chain.

The contrastive analysis of the corpora was carried out with *SketchEngine* software (Kilgariff et al., 2004). The list of metadiscourse markers was based on items identified by previous research (e.g., Hyland, 2005; Lee & Casal, 2014; Ho & Li, 2018) and extended as a result of close reading of sample texts. All concordances were checked in context to ensure that they function as metadiscourse markers. We have decided to prioritise rhetorical functions over lexical realisations. Since the boundaries of metadiscourse categories are fuzzy (cf. McGrath & Kuteeva, 2012), the items were assigned to the predominant metadiscourse category they express, and some items have been assigned to more than one metadiscourse category. For example, *I* functions as self-mention, although it may also be part of phrases functioning as frame markers or personal asides; however, in the case of frame markers it is the verbal or nominal item that defines the rhetorical function of the phrase, and in the case of personal asides their function of parenthetical comment. Therefore, duplicity, if any, is assumed to be minimal and does not substantially impact the results. Careful

contextual analysis was used to explore functional variation within individual metadiscourse categories. The results of the frequency analysis were normalised per 10,000 words (pttw) to allow comparison across all three corpora. The statistical significance of differences was calculated using the non-parametric log-likelihood statistical test (Rayson et al., 2004) with a significance level set at <0.05 (<0.001 is used in tables to indicate very low p-values).

3 Results and discussion

The results of the frequency analysis (Table 2) show that all groups of writers represented in the corpora use more interactional metadiscourse than interactive metadiscourse. This aligns with the findings of several recent studies reporting stronger preference for interactional metadiscourse in university students' academic discourse (e.g., Aull & Lancaster, 2014; Ho & Li, 2018; Qin & Uccelli, 2019). However, Hyland's (2004) study in L2 postgraduate writing found that interactive markers were more frequent than interactional markers, and expert academic discourse seems to show a prevalence of interactive metadiscourse (e.g., Hyland, 1998; Hyland & Jiang, 2018). These differences may stem from corpora composition, discipline, and analytical approach adopted by the researchers.

In the interactional type, hedges are the most prominent category, followed by boosters and engagement markers, while in the interactive type, the markers with highest incidence are evidentials, followed by transitions. Interestingly, Czech authors use more interactive markers than L1 writers, which may reflect differences in academic writing conventions as well as a focus on transitions and citations typical in academic writing courses. Our findings differ from Hyland (2004), as in Hyland's MTs corpus transitions are more prominent than evidentials and engagement markers exceed the frequency of boosters. The difference in transitions may be explained by the inclusion of intersentential connectors in Hyland's study, while the occurrence of fewer boosters may stem from a lower degree of engagement with the reader on the part of Czech students resulting from the merging of L1 and L2 academic conventions.

**METADISOURSE IN L2 MASTER'S THESES:
THE IMPACT OF ACADEMIC CULTURE AND EXPERTISE**

Metadiscourse	MT		BAWE		RA	
	n	pttw	n	pttw	n	pttw
Interactional	28,681	295.8	13,993	295.8	7,985	328.8
<i>Self-mention</i>	1,043	11.0	434	8.8	461	18.9
<i>Hedges</i>	13,729	140.1	6,430	138.5	3,788	156.1
<i>Boosters</i>	5,033	51.3	2,864	61.72	1,222	50.39
<i>Attitude markers</i>	4,325	45.6	1,806	36.7	1,049	43.3
<i>Engagement</i>	4,551	47.8	2,459	50.1	1,465	60.1
Interactive	20,713	218.5	7,601	154.7	4,565	187.9
<i>Transition markers</i>	5,089	53.7	1,726	35.1	1,067	43.9
<i>Frame markers</i>	1,615	17.0	350	7.1	252	10.4
<i>Endophoric markers</i>	3,943	41.6	482	9.8	697	28.7
<i>Evidentials</i>	6,553	69.1	3,805	77.49	1,746	71.8
<i>Code glosses</i>	3,513	37.1	1,238	25.2	803	33.1

Table 2: Metadiscourse across the corpora (pttw)

As Table 3 shows, differences in the overall occurrence of interactional and interactive metadiscourse across the three corpora are significant. The majority of specific metadiscourse categories also yield significant variation. No significant variation has been found in boosters, attitude markers and evidentials across the MT and RA corpora and in engagement across the MT and BAWE corpora. Apart from reflecting genre differences, this may indicate that differences along the expertise dimension are more prominent than those along the culture dimension.

Corpora	MT vs BAWE		MT vs RA		BAWE vs RA	
	<i>LL-G²</i>	p-value	<i>LL-G²</i>	p-value	<i>LL-G²</i>	p-value
Interactional	31.9630	<0.001	147.7864	<0.001	101.2315	<0.001
<i>Self-mention</i>	12.5444	<0.001	88.7657	<0.001	129.3498	<0.001
<i>Hedges</i>	45.4972	<0.001	16.1074	<0.001	72.2415	<0.001
<i>Boosters</i>	16.1128	<0.001	1.1264	0.288	19.3205	<0.001
<i>Attitude markers</i>	61.1989	<0.001	3.2522	0.071	243.1394	<0.001
<i>Engagement</i>	2.8607	0.09	55.5445	<0.001	31.2252	<0.001
Interactive	707.2298	<0.001	655.4316	<0.001	106.77	<0.001
<i>Transition markers</i>	247.5619	<0.001	40.2442	<0.001	32.1230	<0.001
<i>Frame markers</i>	259.7294	<0.001	60.571	<0.001	20.0836	<0.001
<i>Endophoric markers</i>	1285.507	<0.001	89.6620	<0.001	334.1503	<0.001
<i>Evidentials</i>	961.8957	<0.001	2.0636	0.1508	7.9733	0.004
<i>Code glosses</i>	144.1828	<0.001	8.8182	0.0029	35.9615	<0.001

Table 3: Significance of difference in metadiscourse across the corpora (significance level <0.05)

In the following sections we discuss in detail the use of metadiscourse categories across the three corpora.

3.1 Interactional metadiscourse

In terms of frequency, the distribution of interactional metadiscourse categories across the corpora shows a similar tendency: the most prominent category is hedges, followed by boosters in the learner corpora and engagement in the RA, the third position is occupied by engagement in the learner corpora and boosters in the RA, the fourth by attitude markers, while the least frequent category is self-mention. The prominence of hedges and boosters as markers of epistemicity is hardly surprising as they allow writers to modulate the degree of certainty and commitment to their claims, which is essential to academic persuasion, while the importance of engagement resides in its potential to involve readers in the argument and thus persuade them to accept the writer's views and interpretations.

3.1.1 Self-mention

Despite being the least frequent interactional metadiscourse category, self-mention is a highly important marker as it allows writers to indicate their personal stance and gain visibility for themselves and their work.

The frequency of realisation types of pronominal self-mention (Table 4) confirms the prominence of this stance marker in expert discourse. Its occurrence in the RA corpus is significantly higher than in the learner corpora.

Corpora	Self-mention											
	<i>I</i>		<i>we</i>		<i>my</i>		<i>our</i>		<i>me</i>		<i>us</i>	
	n	pttw	n	pttw	n	pttw	n	pttw	n	pttw	n	pttw
MT	644	6.79	56	0.6	281	2.96	4	0.4	57	0.6	0	0
BAWE	375	7.64	0	0	45	0.92	0	0	14	0.2	0	0
RA	317	13.0	21	0.8	97	3.99	6	0.2	17	0.7	0	0

Table 4: Frequency of realisation types of self-mention across the corpora

As evidenced by the lesser incidence of self-mention in the MT and BAWE corpora, backgrounding authorial presence seems to be a distinctive feature of learner discourse motivated by a reluctance to show full commitment to views and claims (cf. Hyland, 2004; Lee & Deakin, 2016; Liu & Zhang, 2022). L2 scholars' English-medium discourse has also been shown to display lesser use of self-mention (e.g., Dontcheva-Navratilova, 2013; Liu & Zhang, 2022; Loréz Sanz, 2011), which may have an impact on Czech graduates' texts.

The less frequent use of pronominal self-mention in Czech students' theses is compensated by a high frequency of occurrence of nominal forms such as abstract rhetors (e.g., *the thesis*, *this chapter*) or the phrase *the author (of this thesis)* (Dontcheva-Navratilova, 2013, 2023a). Self-mention by exclusive *we/our* occurs in five of the 48 theses in the MT corpus; this may be interpreted as interference of the L1 academic writing conventions as such use of exclusive first-person plural pronouns is typical of Czech academic discourse (Čmejrková & Daneš, 1997). No incidence of exclusive *we* has been found in the L1 corpora. While most self-mention markers occur in the agentive subject position, granting a high degree of visibility to the writer (1), in the MT and RA corpora there is also a substantial presence of possessive forms (2), which allow writers to create a strong association between the researcher and their data, findings or interpretations (*my corpus*, *my analysis*, *my view*); since AEs are not research-oriented, they show a low incidence of these patterns.

- (1) *In presenting these findings **I argue** that the description, or reference to, collectives is not the same as enacting those collectives.* (RA_LIT_05)

An analysis of the rhetorical functions of self-mention reveals further differences. The most prominent in all corpora is the researcher role (Dontcheva-Navratilova, 2023a) related to the description of data collection and presentation and interpretation of results (2) (10.9 pttw in RA, 6.7 pttw in MT and 5.6 pttw in BAWE). The importance of the discourse organiser role is greater in the longer texts or RAs (3.5 pttw) and MTs (2.0 pttw), where assisting readers through the text is of primary importance (2).

- (2) *In part four **I present** and discuss **my** findings, drawing a classification of the observed metaphors and comparing them to the aforementioned previous study.* (MT_LIN_04)

The most powerful, authorial roles of arguer (1) and evaluator are particularly prominent in the RA corpus (3.1 and 0.7 pttw respectively), where the researcher steps into the text to put forward claims, comment on findings and evaluate previous research, thus assuming a position of authority and enhancing their visibility.

The reflexive-self role locating the writer in a specific socio-cultural context and casting their autobiographical self into the text (cf. Starfield & Ravelli, 2006) is most frequently used in introductions to MTs, where Czech graduates explain their motivation for choosing a topic or narrate personal stories connected to their studies (3). This may be seen as a self-disclosure strategy for personalising their work and stressing their involvement with the research topic or methodology.

- (3) *My interest in developing learner autonomy was sparked by my first encounter with ALL which happened during a JAPO course [...].* (MT_EDU_10)

These findings indicate that the realisations and functions of self-mention vary along the genre, cultural and expertise dimensions across the corpora. In addition, the reluctance of Czech graduates to employ self-mention more frequently may be induced by the examination context in which the MT is set, as well as by the influence of academic writing style guides and Czech academic writing conventions, where the use of self-mention is generally avoided.

3.1.2 Hedges

Despite being the dominant stance category in all corpora, hedges show significant variation in frequency of occurrence between the RA and the learner corpora, revealing variation along the expertise dimension (Table 3).

A comparison of the realisations of hedges (Table 5) shows that in agreement with previous research (e.g., Hyland, 1998; Mur-Dueñas, 2021; Wu & Paltridge, 2021), adverbs (e.g., *often, probably, usually*), modal verbs (e.g., *may, might, would*) and epistemic lexical verbs (e.g., *claim, suggest, indicate*) are the most frequent realisations of hedges in all corpora.

Hedges realisations	MT		BAWE		RA	
	n	pttw	n	pttw	n	pttw
<i>Adjectives</i>	1,152	11.6	449	9.7	302	12.5
<i>Adverbs</i>	4,452	45.5	1,775	38.3	855	35.3
<i>Modal verbs</i>	4,030	41.1	1,640	35.3	1,158	47.6
<i>Semi-modal verbs</i>	776	8.1	584	11.8	260	10.7
<i>Epist. lexical verbs</i>	2,395	25.2	1,489	30.3	871	35.8
<i>Epist. nouns</i>	314	3.2	334	7.1	207	8.5
<i>Phrases</i>	610	6.2	159	3.4	135	5.5

Table 5: Realisations of hedges across the corpora

However, adverbs (4) are considerably more frequent and epistemic lexical verbs are significantly less frequent in the MT corpus than in the L1 corpora, which indicates cross-cultural variation. Modal verbs (e.g., *may, might, could*) are more prominent in the research-oriented genres than in AEs (35.3 pttw). The use of semi-modal verbs (the verbs *seem* and *appear* characterised by a low semantic load) across the corpora shows the same tendency as displayed by epistemic lexical verbs. The reasons for Czech graduates' preferences may stem from academic writing instruction, which dedicates ample time to the use and

practice of stance adverbials but pays attention only occasionally to the rhetorical potential of epistemic lexical verbs and semi-modals.

- (4) ***Generally**, the students, who expressed positive feelings about the project, **usually** stated something along the lines that they did find it useful even though they did not use its full potential.* (MT_EDU_10)

The less frequent use of epistemic lexical verbs (5) for making claims may also be explained by their typical co-occurrence with first person pronouns (e.g., *I argue, I propose*), which contradicts the advice of instructors and style guides to avoid personality. However, distancing from claims may also be regarded as a strategic choice aimed at preventing criticism on the part of the students.

- (5) *What the recording conveys instead, **I propose**, is the sensation of headlines announcing a feature story or breaking news – a sound effect closer to a newsboy's cry than to a headlined placard.* (RA_LIT_08)

The other realisations of hedges (adjectives, epistemic nouns and phrases) are less frequent. The higher rate of phrasal hedges used by Czech students seems to indicate a reliance on set phrases, such as *in general, in my view, from this perspective*.

An analysis of the functions of hedges reveals further differences. In all corpora, content-oriented hedges conveying the precision and reliability of provided information (e.g., *probable, may, likely*) prevail over participant-oriented hedges (Dontcheva-Navratilova, 2023b), which modulate writer-reader interaction (e.g., *claim, appear, in my view*). However, the ratios of content and participant-oriented hedges differ. The MT corpus displays the most substantial difference between the ratios of the two types of hedges (70.2% vs 29.8%), while in the L1 corpora this difference is smaller (58.8% vs 41.2% in BAWE and 62.2% vs 37.8% in RA) (cf. Lee & Deakin (2016) reporting a similar tendency in L1 vs L2 university student writing). The lower proportion of participant-oriented hedges in the MT corpus might be attributed to power relations in the context of a high-stakes examination. It is likely that Czech graduates use fewer participant-oriented hedges to restrict the dialogic space and protect themselves from possible criticism, but at the same time employ content-oriented hedges to present their claims and positions as uncertain and comment cautiously on the views and claims of others (cf. Aull & Lancaster, 2014; Qiu & Ma, 2019).

3.1.3 Boosters

Overall, boosters show no significant variation across the MT and the RA corpora, but their rate is significantly higher in the BAWE corpus (Tables 2 and 3), probably because without the support of research data writers of AEs feel the need to express their views with a higher degree of commitment and certainty.

The most frequent realisations of boosters across the corpora are adverbs (e.g., *always*, *clearly*), followed by lexical verbs (e.g., *show*, *demonstrate*), adjectives (e.g., *clear*, *obvious*) and phrases (e.g., *in fact*, *no doubt*) (Table 6); the occurrence of modal verbs is insignificant, as it is represented by a single item (*must*).

Boosters realisations	MT		BAWE		RA	
	n	pttw	n	pttw	n	pttw
<i>Adjectives</i>	692	7.1	435	9.4	139	5.7
<i>Adverbs</i>	2,419	24.7	1,320	28.4	572	23.6
<i>Modal verbs</i>	14	0.1	10	0.2	4	0.2
<i>Lexical verbs</i>	1,141	11.6	758	16.3	386	15.9
<i>Phrases</i>	767	7.8	341	7.3	121	4.9

Table 6: Realisations of boosters across the corpora

The BAWE corpus displays the highest rate of adverbs as boosters, while the difference across the RA and MT corpora is not significant, the most frequently used items being *always*, *actually*, *clearly*, *highly*, *indeed* and *particularly*. Similarly to hedges, boosting by epistemic lexical verbs (*demonstrate*, *show*) is more prominent in the L1 corpora, which confirms the existence of cross-cultural variation. The higher frequency of adjectives (6) in the learner corpora confirm that learner writers tend to convey their stance with adjectives and adverbs, while experienced writers are likely to opt for epistemic lexical verbs (Wu & Paltridge, 2021).

- (6) *Therefore it is **clear** that the differences in the way poems and novels organize their stories lies mainly in the narrative structures.* (BAWE_ENG_114)

Out of the two functions performed by boosters, that is, emphatics (*clear*, *in fact*, *show*, *demonstrate*, *by no means*) and amplifiers (*always*, *clearly*, *certainly*), the former show a higher ratio across all corpora (61.5% in MT, 65.6% in BAWE and 64.3% in RA). The marked presence of emphatics in the L1 corpora may reflect an effort to express commitment to views and certainty in results interpretation intended to balance the caution and tentativeness expressed

by hedges. The slightly higher ratio of amplifiers (*always, clearly, definitely*) in the MT corpus may be interpreted as the projection of an overgeneralised and assertive stance reported as characteristic of less advanced writers (Aull & Lancaster, 2014).

Regarding the interplay of hedges and boosters, seen as opposite sides on a certainty scale, our results show that learner writers use considerably fewer hedges and slightly more boosters than expert writers, thus confirming the findings of previous research reporting a tendency on the part of learners to express stronger commitment to views and open a restricted dialogic space for the negotiation of their claims (Qiu & Ma, 2019; Wu & Paltridge, 2021). Yet Czech students' use of hedges and boosters is closer to that of expert writers than L1 learner writers, probably due to their more advanced socialisation in academia, the gradual shift towards English academic writing conventions, and because of their efforts to be cautious and deferent in the context of a high-stake examination.

3.1.4 Attitude markers

Attitude markers, which convey the writers' evaluative assessment of the propositional content, do not show significant variation across the MT and the RA corpora (Table 3), while L1 learners use them significantly less frequently. These findings seem to be attributable mainly to the similarity of the research genres of MTs and RAs. Master's students and expert writers are assumed to be more aware than L1 learners of the necessity of involving their readers in the discussion of the research, acknowledging its limitations and indicating the importance and relevance of their findings.

In all three corpora research-oriented attitude markers (Thetela, 1997) predominate over topic-oriented attitude markers (27.7 vs 17.9 pttw in MT corpus, 29.6 vs 13.7 pttw in RA and 22.3 vs. 14.4 pttw in BAWE). All types of writers choose to use attitude markers primarily to evaluate their own research and its findings (7), while evaluation of previous research by other scholars and future research is relatively rare.

- (7) *A surprising outcome of this analysis, however, was the number of non-equivalent idioms.* (MT_LIN_14)

Attitude is most straightforwardly expressed using attitudinal nouns, adjectives, verbs and adverbs. Given their evaluative potential, adjectives were identified as the most frequent in all three corpora, vastly outnumbering the other three realisation types (Table 7). They constitute approximately 62 per cent of realisations of attitude markers in the MTs (29.7 pttw), 65 per cent in the BAWE

(22.8 pttw), and 63 per cent in the RAs (27.6 pttw). The other three realisations display the same order and a similar frequency of occurrence across the three corpora, nouns ranking second, adverbs third and verbs fourth.

	MT		BAWE		RA	
	n	pttw	n	pttw	n	pttw
<i>Adjectives</i>	2,818	29.7	1,119	22.8	669	27.6
<i>Nouns</i>	748	7.9	316	6.4	192	7.9
<i>Adverbs</i>	417	4.4	191	3.9	97	4.0
<i>Verbs</i>	342	3.6	180	3.6	91	3.7

Table 7: Realisation types of attitude markers across the corpora

The range and variety of evaluative expressions identified as attitude markers was also very similar across the three corpora. In all four categories, the lists of most frequent expressions largely overlap: adjectives – *important, key, interesting, surprising, crucial*; nouns: *importance, difficulty, value, insight*; adverbs – *interestingly, importantly, essentially*; and verbs – *support, contribute*, etc. Overall, the evaluation with attitude markers mainly highlights the importance of the research, notable results, and key methodological justifications.

3.1.5 Engagement

The alignment dimension of engagement is more prominent in the RA corpus than in the learner corpora (Table 8). The engagement resources comprise reader reference, personal asides, questions, directives and appeals to shared knowledge. Due to their rare occurrence and frequent overlap with other categories, such as self-mention and attitude markers, personal asides are not analysed here.

Engagement	MT		BAWE		RA	
	n	pttw	n	pttw	n	pttw
<i>Reader reference</i>	1,489	15.7	1,087	22.1	650	26.7
<i>Shared knowledge</i>	1,770	18.6	842	17.0	418	17.1
<i>Directives</i>	1,205	12.6	393	8.0	271	11.2
<i>Questions</i>	87	0.9	137	2.8	126	5.1

Table 8: Frequency of engagement sub-categories across the corpora

As shown in Table 8, the most frequent engagement markers in all corpora are reader reference and appeals to shared knowledge. However, while difference in shared knowledge markers is not significant, reader reference is considerably

more prominent in the L1 corpora, indicating cross-cultural variation. The main function of reader reference, typically realised by the inclusive *we*, is to build proximity with readers by presenting them as belonging to the same 'in-group' as the writer and thus as sharing the same values and views. In the RA corpus, the use of an inclusive *we* typically represents the reader as a co-researcher who shares disciplinary common ground with the writer and follows the argument chain to reach suggested conclusions (8). In the learner corpora, however, the group with which the reader is invited to identify is fuzzier and may often be interpreted as people in general.

- (8) *When we choose to engage with how accounts are constructed as well as what interviewees seem to be saying, we come to understand more clearly how interviewers and interviewees assemble particular discursive resources in co-constructing clarity and seeming reliability too.* (RA_LIN_09)

Appeals to shared knowledge are instrumental in enticing the reader to accept the interpretations, positions and views of the writer, which most likely explains the lack of significant variation across the corpora. The most frequent realisations of appeals to shared knowledge in all corpora are *obvious*, *of course*, *evidently* and *apparently*. There is, however, a slight difference in the placement of these markers, with Czech graduates tending to locate them in sentence initial position (9). Apart from indicating the possible influence of instruction in academic writing, this might stem from the greater visibility of shared knowledge markers in initial position, which helps learners notice and subsequently use them (Dontcheva-Navratilova, 2023b).

- (9) *Of course, facing the reader with seemingly meaningless passages as well as with their authoritative-sounding counterparts is not the only language-related technique Burroughs employs in the text.* (MT_LIT_10)

The use of directives seems to be impacted by genre variation, as their frequency is similar in the research genres represented in the MT and RA corpora but substantially lower in BAWE. Out of the three possible realisations of directives (i.e., imperatives, obligation modals and predicative adjectives) obligation modals are the preferred choice in all corpora (10). Imperatives (e.g., *consider*, *see*) and predicative adjectives (e.g., *it is necessary to*, *it is important to*) are more prominent in the MT and RA corpora, as writers seem to be more prone to assuming a position of authority (Hyland & Jiang, 2019), probably based on their knowledge and research results and despite the imposition and face-threat that these forms imply.

- (10) *Students and teachers **should** be trained in how to master pronunciation more effectively, and tools such as the phonemic chart and on-line courses **ought to** be readily available to the teaching staff.* (BAWE_LIN_152)

The functions of directives (textual, cognitive and physical acts) also differ across the corpora. While, as is typical of most soft disciplines, physical acts are practically absent, textual acts (realised mostly by *see*) have a noticeable presence in the MT (23%) and the RA (32.5%) corpora, albeit only a minor one in the BAWE (5%). This results from the length of the text and the frequent use of tables, paragraphs, examples and appendices in RAs and MTs, which require writers to make cross-references to these text components. The predominant function of directives in all corpora is to perform cognitive acts which strive to direct the readers' attention to important points and guide them towards intended interpretations (e.g., *it is important to note, consider key characteristics of the data, contextual factors must be acknowledged*).

The rhetorical potential of questions to focus the attention of the reader on key points in the argument is exploited primarily by expert writers. In the learner corpora, especially in the MT corpus, they occur relatively rarely. This suggests that the mastering of this explicitly dialogic feature is associated with rhetorical maturity and expertise.

3.2 Interactive metadiscourse

The significantly higher frequency of interactive markers in the MT corpus in comparison with the L1 corpora results primarily from the high rate of transitions, endophoric markers, frame markers and code glosses (Tables 2 and 3). The most frequent interactive marker across all corpora are evidentials, which confirms the key role of reference to previous research in academic discourse. However, they are the only interactive marker that does not show significant variation across the RA and the learner corpora. Transition markers are second in frequency “as they represent writers' attempts to ensure readers are able to correctly recover their intentions” (Hyland, 2004, p. 140).

3.2.1 Transition markers

Transition markers show significant differences in frequency across the three corpora: they are the most frequent in the MT corpus and the least frequent in RAs. This seems to reflect the ability of expert writers to indicate logical relations in various ways, as well as the strong emphasis on transitions in academic writing instruction, which often leads to overuse of surface cohesive means (cf. Dontcheva-Navratilova et al., 2020). The three sub-categories of transitions

reflecting the basic types of relations between the ideas or arguments proposed, namely ‘addition’ (e.g., *moreover*, *in addition*), ‘comparison’ pointing either to similarity (e.g., *similarly*, *likewise*) or difference between arguments (e.g., *but*, *in contrast*, *however*), and ‘consequence’ (e.g., *thus*, *therefore*), also show variation across the corpora (Table 9).

Transitions	MT		BAWE		RA	
	n	pttw	n	pttw	n	pttw
<i>Addition</i>	1,513	15.9	481	9.8	180	7.4
<i>Comparison</i>	2,251	23.7	1,496	30.4	491	20.9
<i>Consequence</i>	1,525	16.1	511	10.4	205	8.4

Table 9: Frequency of transitions sub-categories across the corpora

While in the L1 corpora addition and consequence are not very frequent as these relations can be implied, they are prominent in the MT corpus showing a tendency toward overexplicitness (11). The most prominent sub-category of transitions in the L1 corpora is comparison (12), which is particularly high in the BAWE corpus probably because of the comparative American studies essays.

- (11) ***In addition**, as simulation games promote creative and imaginative thinking, they have positive effects on the development of these capacities* (Kusnierek, 2015). ***Furthermore**, as it has been suggested, participation in simulations requires a great deal of autonomy on the part of the student [...].* (MT_EDU_14)
- (12) ***Likewise**, Eliot uses Colonel Townley’s status as an outsider, as a vehicle to introduce the reader to the novel.* (BAWE_ENG_075)

The lists of the five most frequent transition markers in the three corpora display a certain degree of similarity. While the markers of contrast *however* and consequence *therefore* are highly prominent in all corpora, the top five also include *thus* and the markers of addition *moreover* and *furthermore* (11) in MTs; *thus* and the conjunctions *but* and *and* in RAs; and *yet*, *on the other hand* and *furthermore* in the BAWE corpus. While expert writers tend to use conjunctions performing a metadiscoursal function rather frequently, learner writers rely more heavily on adverbs to express transitions.

3.2.2 Frame markers

Frame markers exhibited the highest frequency in the MTs (17.0 pttw), compared to RAs (10.4 pttw) and AEs (7.1 pttw). This result confirms that “longer papers, of course, require more explicit structuring to ensure readers are able to follow the direction of the argument” (Hyland & Jiang, 2018, p. 16).

However, fewer frame markers in expert writing may indicate a higher level of sophistication in text organisation (Noble, 2010, p. 160).

A comparison of the four frame marker subcategories (Hyland, 2005, p. 51) shows varying frequencies across the corpora (Table 10).

Frame markers	MT		BAWE		RA	
	n	pttw	n	pttw	n	pttw
<i>Sequencing</i>	749	7.9	64.0	1.3	54	2.2
<i>Label stages</i>	130	1.4	85.0	1.7	29	1.2
<i>Announcing goal</i>	658	6.9	114.0	2.3	121	5.0
<i>Topic shift</i>	78	0.8	87.0	1.8	48	2.0
TOTAL:	1,615	17.0	350.0	7.1	252	10.4

Table 10: Distribution of subcategories of frame markers across corpora

The occurrence of sequencers in the MT corpus was significantly higher than in the L1 corpora (13), yet common devices such as *first*, *then*, and *finally* were identified across all corpora. Stage labellers were rare in all three corpora, though summarising labellers (e.g., *in conclusion*, *to sum up*, *to summarise*) were most prevalent in the learner corpora, indicating that novice writers tend to use fixed phrases and an impersonal tone in summarising their points.

- (13) *To start with*, compliments are used to ease the process of communication, so by their very nature they are exchanged by the participants of a conversation. (MT_LIN_12)

Goal announcers (e.g., *aim*, *focus*, *intend to*, *objective*, *seek to*) were more common in the MT corpus than in the shorter L1 texts. Self-mention co-occurring with verbal phrases to announce goals, such as *I/we want to*, *I argue*, *I intend*, were primarily used by expert writers, while Czech graduates preferred an impersonal tone. Finally, topic shifters (e.g., *in regard to*, *move on*, *now*, *shift to*, *so*) occurred more frequently in the L1 corpora. The most common topic shifter in RAs was *so* (0.9 pttw), compared to BAWE (0.2 pttw) and MT (0.04 pttw). The absence of *so* in Master's theses is unsurprising, given that it is not regarded as a conventional topic shifter in academic writing.

3.2.3 Endophoric markers

The frequency of endophoric markers (e.g., see Figure 2, as noted above) varies significantly across the three corpora (Tables 2 and 3). Their occurrence is very high in the MT corpus due to text length, where guiding readers to specific

text sections and spotlighting (e.g., examples, visual aids, and research findings) is highly important (Hyland, 2005), and low in the BAWE corpus comprising shorter texts with less need for cross-referencing.

Table 11 summarizes the occurrence of the three sub-categories of endophoric markers (anaphoric, cataphoric, and non-directional) across the corpora.

Endophoric markers	MT		BAWE		RA	
	n	pttw	n	pttw	n	pttw
<i>Anaphoric</i>	1,548	16.3	281	5.7	246	10.7
<i>Cataphoric</i>	1,273	13.4	130	2.6	131	5.4
<i>Non-directional</i>	1,122	11.8	71	1.4	320	13.2
TOTAL:	3,943	41.6	482	9.8	697	28.7

Table 11: Frequency of endophoric markers across the corpora

In each corpus, anaphoric reference predominates over cataphoric, as authors often return to previously mentioned points. Anaphoric reference is most frequent in learner writing (14), while experts favour non-directional (15) and anaphoric reference. The low frequency of non-directional reference in BAWE is due to the shorter length and less formal structure of argumentative essays, resulting in no reference to chapters or sections, and less frequent reference to items incorporated into the text, for example, figures, or tables.

- (14) *As it has been suggested **above**, the main advantage of the model as it is understood today is its relative simplicity and universality of usage.* (MT_EDU_14)
- (15) ***Table 6** summarizes how the different referent types were introduced by the three native-language groups.* (RA_EDU_10)

3.2.4 Evidentials

Evidentials are the most frequent interactive metadiscourse marker across the three corpora. The highest rate of evidentials is present in the BAWE corpus (77.49 pttw), which seems to result from the limited size of the text in which authors need to ground their argument in previous research. The MT and RA corpora do not show significant difference in the frequency of evidentials (69.1 vs 71.9 pttw; $p=0.1508$). This frequency is higher than the rate of 64.1 in post-graduate theses across several soft and hard sciences disciplines as reported by Hyland (2004), which might reflect the choice of disciplines represented in the corpora but also shows that Czech graduates seem to be aware of the importance of citations in academic discourse.

A comparison of the ratio of integral vs non-integral citations (cf. Swales, 1990) shows that the realisation types of citations in the MT corpus differ from those in the L1 corpora (Table 12). While 60.7 per cent of all citations used by Czech graduates are integral, L1 writers show a preference for non-integral forms. A predominance of integral citations in student writing in the social sciences and humanities has been reported by previous research (e.g., Ādel & Garretson, 2006). Professional academic discourse, however, is marked by a prevalence of non-integral forms (cf. Hyland & Jiang, 2017). Thus, the higher prominence of non-integral citations in the RA corpus (64.2%) in comparison to the BAWE corpus (53.6%) and the MT corpus (39.3%) suggests that variation across the corpora is affected by degree of expertise rather than by culture.

Interactive metadiscourse	MT		BAWE		RA	
	n	%	n	%	n	%
<i>Integral citations</i>	3,976	60.7	1,765	46.4	626	35.8
<i>Non-integral citations</i>	2,577	39.3	2,040	53.6	1,120	64.2
<i>Evidentials</i>	6,553	100.0	3,805	100.0	1,746	100.0

Table 12: Integral vs non-integral citations across the corpora (in per cent)

Non-integral citations emphasise the content of the message, as reference to the cited researcher is typically confined to the name of the cited author in parenthesis (16). This allows the writer to summarise large amounts of information and display familiarity with numerous sources without interrupting the flow of the argument.

- (16) *Disagreement tends to revolve around topics or ideas (Grimshaw, 1990b), while disaffiliation or disalignment occurs regarding participants (Kjaerbeck, 2008).* (RA_LIN_02)

By including the cited author's name as a clause element, integral citations give high prominence to the cited researcher, who is often positioned as an authority supporting the author's view (17), and they help the writer highlight selected source content and evaluate reported material. Integral citations may take the form of paraphrase, which according to Shi (2010) students perceive as enhancing the academic quality of their MT, or a direct quote, which gives prominence to the original wording of quoted previous research.

- (17) *For instance, Scrivener (2011) pointed out that simulations are in fact only "large scale role-plays" (p. 224).* (MT_EDU_14)

Set phrases including the name of the cited author functioning as an adjunct are significantly more frequent in the MT corpus than in the L1 corpora (e.g., *according to NAME* and *as NAME+VERB/VERB+NAME*). This suggests that Czech graduates tend to rely on academic phrases acquired in academic writing courses to enhance the accuracy and the academic style of their texts but lack sufficient expertise in the use of evidentials. To improve intertextual referencing, students need to invest more time in developing their writing proficiency and familiarising themselves with disciplinary and genre conventions (e.g., Thompson, 2005).

3.2.5 Code glosses

The frequency of code glosses (e.g., *i.e.*, *that is*, *in other words*) shows significant variation across the corpora (Table 2). Czech writers reformulated and clarified their statements most frequently (37.1 ptw); in comparison, L1 student writers used code glosses considerably less frequently (25.2 ptw), with L1 expert writers being in the middle (33.1 ptw). A closer look at the two functions of code glosses, reformulation and exemplification (Hyland, 2007), reveals that in all three corpora, exemplification predominates over reformulation, which shows that both learner and expert writers recognize its importance in academic argumentation. Since the distribution of exemplification markers is very similar across the three corpora (Table 13), the greatest difference can be found in the use of reformulation.

Functions of code glosses	MT		BAWE		RA	
	ptw	%	ptw	%	ptw	%
<i>Reformulation</i>	16.3	44.0	5.8	23.0	13.5	40.8
<i>Exemplification</i>	20.8	56.0	19.4	77.0	19.6	59.2
Total	37.1	100	25.2	100	33.1	100

Table 13: Subfunctions of code glosses across the three corpora

The highest incidence of reformulation markers is found in the MTs. This may be attributable to the character of the genre, which requires that authors are able to demonstrate knowledge and understanding of the theories, methods and terminology of a given discipline, and more specifically of the research problem studied (18). On the other hand, AEs in the BAWE corpus are considerably shorter and do not provide so many opportunities for rephrasing or explanations. MTs and RAs proved similar since they are both research-process genres and their authors often use code glosses to explain, define or clarify their statements.

- (18) *'Synonymy' is a type of paradigmatic relations, i.e. relations that "reflect the semantic choices available at a particular structure point in a sentence" (Cruse 2000: 148), and is generally defined as sameness or identity of meaning. (MT_LIN_12)*

Specific code glosses the writers opt for seem to reflect cultural differences reflecting different academic writing conventions. When reformulating, Czech students overwhelmingly prefer the abbreviation *i.e.*, which accounts for 31 per cent of all reformulation markers in the MT corpus. Both expert and learner L1 writers employ more varied devices, such as *that is*, *in other words*, *mean*, *specifically*. Interestingly, Murillo (2018) discovered that *i.e.* was the most commonly used reformulation marker in English-medium RAs by Czech authors. This suggests that Czech novice and expert writers both rely heavily on this simple, grammaticalised form which does not pose any problems in text production and comprehension.

4 Conclusion

This article has studied metadiscourse in MTs by L2 (Czech) graduates to explore how L2 learners organise their discourse, express an evaluative stance and engage in a dialogue with readers. Seeking to contribute to intercultural rhetoric studies, it has also carried out a contrastive analysis between the use of metadiscourse in three corpora representing Czech students' MTs and L1 university students and expert writers' academic discourse. The findings indicate that the use of metadiscourse in Czech graduates' MTs is influenced by several interrelated factors, the most important of which seems to be the level of expertise; academic writing culture (L1) and genre appear to affect realisation choices and functional specialisations of specific categories of metadiscourse markers.

While in all corpora interactional metadiscourse is more prominent than interactive metadiscourse, the realisation patterns and preferences for specific functions of metadiscourse markers of the three groups of authors vary significantly. Czech graduates use fewer interactional metadiscourse markers than do expert writers, which may stem from a reluctance to display a high degree of authorial visibility and to engage overtly with the reader. This may be explained by cultural differences (i.e., differences in academic writing conventions) as well as by the examination context in which the MT is set, as students are likely to opt for deference, humility and impersonality, so seeking to avoid criticism and meet the expectations of the examiners, who are the primary readers of the theses. Czech students' frequent use of interactive metadiscourse reflects their effort to structure and enhance their academic style using set phrases to achieve text coherence.

Within interactional metadiscourse categories, the more prominent occurrence of self-mention, hedges and engagement markers in the published RAs in comparison to the learner corpora indicate variation along the expertise dimension. Expert writers seek recognition within their discourse community by striving to enhance their visibility, creating a rapport with readers and opening a dialogic space for negotiating suggested interpretations. This is also reflected in the more frequent occurrence of the most powerful self-mention roles of arguer and evaluator in expert writers' texts. In contrast, learner writers are less likely to step into the text by using self-mention and to create an in-group relationship with their readers (i.e., examiners), by employing engagement markers. Thus, in consonance with findings of previous research into L2 learner discourse (Hyland, 2004; Qiu & Ma, 2019; Wu & Paltridge, 2021), Czech graduates use fewer hedges and slightly more boosters than L1 expert writers, which makes their texts look more assertive than is typical of academic discourse. The lesser occurrence of self-mention can also be impacted by cultural differences, particularly the impact of Czech academic writing norms that discourage the use of personality.

Variation in interactive metadiscourse is clearly influenced by text length and genre. The most prominent interactive markers across all corpora are evidentials and transition markers. The very high frequency of evidentials in AEs is somewhat surprising, but it seems to result from the need to anchor the argument in previous research within a very short text, which results in a high relative frequency of citations. Czech graduates use more integral citations than non-integral citations, while in the L1 corpora the situation is reversed. This may be explained by a lower degree of rhetorical maturity on the part of the students, but also by the belief that the use of integral citations by paraphrase enhances the academic quality of an MT (Shi, 2010). Transition markers and code glosses are most prominent in MTs, followed by RAs indicating the need to make explicit the development of the argumentation chain in research genres. The need to mark discourse organisation and guide the reader through the extensive texts of MTs and RAs explains the more frequent occurrence of endophoric and frame markers in these genres. In the learner corpora these are realised primarily by a restricted range of fixed phrases, while expert writers use a wider repertoire of metadiscourse resources.

In conclusion, the findings allow us to draw a few pedagogical implications. In agreement with Hyland (2004, p. 148), we argue that "conscience raising is crucial in L2 academic writing instruction and for teachers this means helping students to move beyond the conservative prescriptions of the style guides"; this may be achieved by encouraging students to analyse their own writing and by providing them with expert models for comparison. Special attention should

be paid to making students aware of intercultural differences and familiarising themselves with disciplinary and genre conventions (Thompson, 2005). Explicit instruction on metadiscourse features in academic writing courses tailored to the needs of graduates (Lee & Deakin, 2016; Wingate, 2012) could assist them in using these rhetorical features strategically. Such courses would ideally combine peer and teacher feedback allowing students to reflect on how they use metadiscourse devices to express their stance, engage with readers and organise their texts. As Lee and Deakin (2016, p. 32) argue, “making these pervasive yet “hidden” dimensions of persuasive writing explicitly visible” could enable students to make conscious rhetorical choices in their efforts to enhance the persuasiveness of their discourse.

Finally, this study has its limitations, as it has focused on only one L2 context and a limited set of disciplines in the soft sciences. Therefore, our results should not be overgeneralised; they should be verified by future research exploring metadiscourse in a wider range of cultural contexts and disciplines and combining the ‘marker’ and ‘move’ approaches to the study of metadiscourse (Ädel, 2023). Further research should also consider the effect of AI tools on MA thesis writing since it may be predicted that they will influence the patterns of use of metadiscourse devices in the final texts.

Acknowledgement

This work was supported by the Czech Science Foundation grant 21-12150S Intercultural variation in writer-reader interaction in English-medium academic discourse by Czech and Anglophone novice writers.

References

- Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. John Benjamins.
- Ädel, A. (2023). Adopting a ‘move’ rather than a ‘marker’ approach to metadiscourse: A taxonomy for spoken student presentations. *English for Specific Purposes*, 69, 4–18. <https://doi.org/10.1016/j.esp.2022.09.001>
- Ädel, A., & Garretson, G. (2006). Citation practices across disciplines: The case of proficient student writing. In M. C. Pérez-Llantada, R. Pló Alastrué, & C. P. Neumann (Eds.), *Academic and professional communication in the 21st century: Genres and rhetoric in the construction of disciplinary knowledge* (pp. 271–280). Prensas Universitarias.
- 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/10.1177/0741088314527055>

- Connor, U., & Moreno, A. (2005). Tertium comparationis: A vital component in contrastive rhetoric research. In P. Bruthiaux, D. Atkinson, W. Eggington, W. Grabe, & V. Ramanathan (Eds.), *Directions in applied linguistics: Essays in honor of Robert B. Kaplan* (pp. 153–164). Multilingual Matters.
- Crosthwaite, P., Cheung, L., & Jiang, F. (2017). Writing with attitude: Stance expression in learner and professional dentistry research reports. *English for Specific Purposes*, 46, 107–123. <https://doi.org/10.1016/j.esp.2017.02.001>
- Čmejrková, S., & Daneš, F. (1997). Academic writing and cultural identity. The case of Czech academic writing. In A. Duszak (Ed.), *Culture and styles of academic discourse* (pp. 40–62). Mouton de Gruyter.
- Dontcheva-Navratilova, O. (2013). Authorial presence in academic discourse: Functions of author-reference pronouns. *Linguistica Pragmensia*, 23(1), 9–30.
- Dontcheva-Navratilova, O. (2023a). Self-mention in L2 (Czech) learner academic discourse: Realisations, functions and distribution across master's theses. *Journal of English for Academic Purposes*, 64, 101272. <https://doi.org/10.1016/j.jeap.2023.101272>
- Dontcheva-Navratilova, O. (2023b). Writer-reader interaction in L2 learner academic discourse: Reader engagement in Czech students' Master's theses. *Linguistica Pragmensia*, 33(2), 117–135.
- Dontcheva-Navratilova, O., Jančaříková, R., Hůlková, I., & Schmied, J. (2020). Theme choices in Czech university students' English-medium Master's theses. *Lingua*, 243, 102892. <https://doi.org/10.1016/j.lingua.2020.102892>
- Duff, P. A. (2010). Language socialization into academic discourse communities. *Annual Review of Applied Linguistics*, 30, 169–192. <https://doi.org/10.1017/S0267190510000048>
- Flowerdew, L. (2015). Using corpus-based research and online academic corpora to inform writing of the discussion section of a thesis. *Journal of English for Academic Purposes*, 20, 58–68. <https://doi.org/10.1016/j.jeap.2015.06.001>
- Guziurová, T. (2022). Glossing an argument: Reformulation and exemplification in L2 Master's theses. *Topics in Linguistics*, 23(2), 18–35. <https://doi.org/10.2478/topling-2022-0009>
- Ho, V., & Li, C. (2018). The use of metadiscourse and persuasion: An analysis of first year university students timed argumentative essays. *Journal of English for Academic Purposes*, 33, 53–68. <https://doi.org/10.1016/j.jeap.2018.02.001>
- Hong, H., & Cao, F. (2014). Interactional metadiscourse in young EFL learner writing: A corpus-based study. *International Journal of Corpus Linguistics*, 19(2), 201–224.
- Hyland, K. (1998). Boosting, hedging and the negotiation of academic knowledge. *Text*, 18(3), 349–82. <https://doi.org/10.1515/text.1.1998.18.3.349>
- Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing*, 13(2), 133–151. <https://doi.org/10.1016/j.jslw.2004.02.001>
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.
- Hyland, K. (2007). Applying a gloss: Exemplifying and reformulating in academic discourse. *Applied Linguistics*, 28(2), 266–285.
- Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, 25(2), 156–177. <https://doi.org/10.1093/applin/25.2.156>
- Hyland, K., & Jiang, K. (2017). Points of reference: Changing patterns of academic citation. *Applied Linguistics*, 40(1), 64–85.

- Hyland, K., & Jiang, F. (2018) "In this paper we suggest": Changing patterns of disciplinary metadiscourse. *English for Specific Purposes*, 51, 18–30. <https://doi.org/10.1016/j.esp.2018.02.001>.
- Hyland, K., & Jiang, K. (2019). *Academic discourse and global publishing: Disciplinary persuasion in changing times*. Routledge.
- Jančáříková, R. (2023). Attitude markers in L2 learners' academic writing: A case study of Master's theses by Czech students compared to L1 students' writings. *Brno Studies in English*, 49(1), 5–31. <https://doi.org/10.5817/BSE2023-1-1>
- Kilgariff, A., Rychlý, P., Smrz, P., & Tugwell, D. (2004). The Sketch Engine. In *Proceedings of the Eleventh EURALEX International Congress* (pp. 105–116), Lorient, France.
- Lahodová Vališová, M. (2024). Endophoric signposting: A contrastive study of textual references in L2 Czech Master's theses and native English academic writing. *Miscelánea: A Journal of English and American Studies*, 70, 15–40. https://doi.org/10.26754/ojs_misc/mj.20249976
- Lee, J. J., & Casal, J. E. (2014). Metadiscourse in results and discussion chapters: A cross-linguistic analysis of English and Spanish thesis writers in engineering. *System*, 46, 39–54. <http://doi.org/10.1016/j.system.2014.07.009>
- Lee, J. J., & Deakin, L. (2016). Interactions in L1 and L2 undergraduate student writing: Interactional metadiscourse in successful and less-successful argumentative essays. *Journal of Second Language Writing*, 33, 21–34. <http://doi.org/10.1016/j.jslw.2016.06.004>
- Liu, G., & Zhang, J. (2022). Interactional metadiscourse and author identity construction in academic theses. *Journal of Language Teaching and Research*, 13(6), 1313–1323. <https://doi.org/10.17507/jltr.1306.20>
- Lorés Sanz, R. (2011). *The study of authorial voice: Using a Spanish–English corpus to explore linguistic transference*. *Corpora*, 6(1), 1–24. <https://doi.org/10.3366/cor.2011.0002>
- McGrath, L., & Kuteeva, M. (2012). Stance and engagement in pure mathematics research articles: Linking discourse features to disciplinary practices. *English for Specific Purposes*, 31(3), 161–173. <https://doi.org/10.1016/j.esp.2011.11.002>
- Mauranen, A. (1993). *Cultural differences in academic rhetoric: A textlinguistic study*. Peter Lang.
- Mur-Dueñas, P. (2011). An intercultural analysis of metadiscourse features in research articles written in English and in Spanish. *Journal of Pragmatics*, 43(12), 3068–3079. <https://doi.org/10.1016/j.pragma.2011.05.002>
- Murillo, S. (2018). Not the same, but how different? Comparing the use of reformulation markers in ELF and in ENL research articles. In P. Mur-Dueñas & J. Šinkūnienė (Eds.), *Intercultural perspectives on research writing* (pp. 237–253). John Benjamins.
- Noble, W. (2010). Understanding metadiscoursal use: Lessons from a 'local' corpus of learner academic writing. *Nordic Journal of Applied Linguistics*, 9(2), 145–169.
- Paltridge, B. (2002). Thesis and dissertation writing: An examination of published advice and actual practice. *English for Specific Purposes*, 21(2), 125–143.
- Pérez-Llantada, C. (2013). 'Glocal' rhetorical practices in academic writing. *European Journal of English Studies*, 17(3), 251–268. <https://doi.org/10.1080/13825577.2013.867180>
- Povolná, R. (2012). Causal and contrastive discourse markers in novice academic writing. *Brno Studies in English*, 38(2), 131–148.

- Qin, W., & Uccelli, P. (2019). Metadiscourse: Variation across communicative contexts. *Journal of Pragmatics*, 139, 22–39. <https://doi.org/10.1016/j.pragma.2018.10.004>
- Qiu, X., & Ma, X. (2019). Disciplinary enculturation and authorial stance: Comparison of stance features among master's dissertations, doctoral theses, and research articles. *Iberica*, 38, 327–348.
- Rayson, P., Berridge, D., & Francis, B. (2004). Extending the Cochran rule for the comparison of word frequencies between corpora. In G. Purnelle, C. Fairon, & A. Dister (Eds.), *Le Poids des Mots: Proceedings of the 7th International Conference on Statistical analysis of Textual Data (JADT 2004)* (pp. 926–936). Presses universitaires de Louvain.
- Sancho-Guinda, C., & Hyland, K. (2012). Introduction: A context-sensitive approach to stance and voice. In K. Hyland & C. Sancho-Guinda (Eds.), *Stance and voice in written academic genres* (pp. 1–14). Palgrave Macmillan.
- Shaw, P. (2003). Evaluation and promotion across languages. *Journal of English for Academic Purposes*, 2(4), 343–357. [https://doi.org/10.1016/S1475-1585\(03\)00050-X](https://doi.org/10.1016/S1475-1585(03)00050-X)
- Shi, L. (2010). Textual appropriation and citing behaviors of university undergraduates. *Applied Linguistics*, 31, 1–24.
- Starfield, S., & Ravelli, J. (2006). The writing of this thesis was a process that I could not explore with the positivistic detachment of the classical sociologist”: Self and structure in New Humanities research theses. *Journal of English for Academic Purposes*, 5(3), 222–243. <https://doi.org/10.1016/j.jeap.2006.07.004>
- Swales, J. (1990). *Genre analysis. English in academic and research settings*. Cambridge University Press.
- Thetela, P. (1997). Evaluated entities and parameters of value in academic research articles. *English for Specific Purposes*, 16(2), 101–118. [https://doi.org/10.1016/S0889-4906\(96\)00022-1](https://doi.org/10.1016/S0889-4906(96)00022-1)
- Thompson, P. (2005). Points of focus and position: Intertextual reference in PhD theses. *Journal of English for Academic Purposes*, 4(4), 307–323. <https://doi.org/10.1016/j.jeap.2005.07.006>
- Thompson, P. (2013). Thesis and dissertation writing. In B. Paltridge & S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 283–299). Wiley-Blackwell. <https://doi.org/10.1002/9781118339855.ch15>
- Thompson, G., & Thetela, P. (1995). The sound of one hand clapping: The management of interaction in written discourse. *Text*, 15(1), 103–127. <https://doi.org/10.1515/text.1.1995.15.1.103>
- Vogel, R. (2008). Sentence linkers in essays and papers by native vs. non-native writers. *Discourse and Interaction*, 1(2), 119–126.
- Wingate, U. (2012). ‘Argument!’ helping students understand what essay writing is about. *Journal of English for Academic Purposes*, 11(2), 145–154. <https://doi.org/10.1016/j.jeap.2011.11.001>
- Wu, B., & Paltridge, B. (2021). Stance expressions in academic writing: A corpus-based comparison of Chinese students’ MA dissertations and PhD theses. *Lingua*, 253, 103071. <https://doi.org/10.1016/j.lingua.2021.103071>

Olga Dontcheva-Navratilova is Associate Professor at the Department of English Language and Literature, Masaryk University, Brno, Czech Republic. She specialises in discourse analysis, genre analysis, stylistics and pragmatics, focusing on academic and political discourse. She is author of *Analysing Genre: The Colony Text of UNESCO Resolutions* (2009) and *Coherence in Political Speeches* (2011) and co-author of *Persuasion in Specialised Discourse* (2020). Her recent work exploring English for academic and specific purposes has been published in a number of international journals and edited volumes.

ORCID ID: <https://orcid.org/0000-0002-0378-7975>

Address: Olga Dontcheva-Navratilova, Department of English Language and Literature, Faculty of Education, Poříčí 9, 603 00 Brno, Czech Republic. [e-mail: navratilova@ped.muni.cz]

Tereza Guziurová is Assistant Professor of English Linguistics at the Faculty of Education, University of Ostrava, Czech Republic. She specializes in discourse analysis and genre analysis. Her research interests include metadiscourse in academic genres and the aspects of writer-reader interaction in academic texts in general. Recently, her research has focused on English as a Lingua Franca and learner language. In 2018 she published a monograph *Metadiscourse in Undergraduate Textbooks and Research Articles in Linguistics*.

ORCID ID: <https://orcid.org/0000-0002-7739-8980>

Address: Tereza Guziurová, Department of English Language and Literature, Faculty of Education, Poříčí 9, 603 00 Brno, Czech Republic. [e-mail: 246275@mail.muni.cz]

Renata Jančaříková is Assistant Professor at the Faculty of Education, Masaryk University in Brno, Czech Republic. Her research interests lie mainly in the area of discourse analysis with a focus on academic discourse and student writing at university level, and British crime discourse and its social aspects. In addition to a number of journal articles, she is the author of *Victims vs Killers in the British Press: Naming Strategies in Murder Reports* (2014) and co-author of the volume *Coherence and Cohesion in English Discourse* (2012). She is co-editor of the journal *Discourse and Interaction*.

ORCID ID: <https://orcid.org/0000-0001-7260-8662>

Address: Renata Jančaříková, Department of English Language and Literature, Faculty of Education, Poříčí 9, 603 00 Brno, Czech Republic. [e-mail: jancarikova@ped.muni.cz]

Marie Lahodová Vališová is Assistant Professor at the Masaryk University Language Centre, Faculty of Medicine in Brno, Czech Republic. She earned her Ph.D. in 2024 with a dissertation on requests and apologies formulated by Czech students in English (EFL). Her research interests include pragmatics, focusing on politeness and speech acts, academic writing with an emphasis on metadiscourse, and ESP course design.

ORCID ID: <https://orcid.org/0000-0001-6411-0012>

Address: Marie Lahodová Vališová, Language Centre Division at the Faculty of Medicine, Masaryk University, Kamenice 553/5, 625 00 Brno, Czech Republic.
[e-mail: marie.lahodova@med.muni.cz]