ENDOPHORIC SIGNPOSTING: A CONTRASTIVE STUDY OF TEXTUAL REFERENCES IN L2 CZECH MASTER'S THESES AND NATIVE ENGLISH ACADEMIC WRITING

MARCADORES ENDOFÓRICOS: UN ESTUDIO CONTRASTIVO DE LAS REFERENCIAS TEXTUALES EN EL INGLÉS ACADÉMICO DE TRABAJOS FIN DE MÁSTER ESCRITOS POR ESTUDIANTES CHECOS EN INGLÉS COMO SEGUNDA LENGUA Y POR HABLANTES NATIVOS

MARIE LAHODOVÁ VALIŠOVÁ
Masaryk University, Czech Republic
marie.lahodova@med.muni.cz
<https://orcid.org/0000-0001-6411-0012>

Abstract

The aim of this study is to contribute to cross-cultural and cross-disciplinary discourse analysis, shedding light on English L2 learners' metadiscursive practices. Focusing on a specialised learner corpus of English-medium Master's theses written by Czech university students, the research explores the occurrence of endophoric markers and their characteristics. To enable cross-linguistic and cross-cultural comparison, a reference corpus of representative L1 English academic discourse was compiled. A new taxonomy used here, which draws on Hyland's list of endophoric markers (2005) and an extensive literature review, categorises endophoric markers into three groups: purely directional, markers using specific words, and a category combining the first two. The study investigates the occurrence and function of these markers, focusing on anaphoric, cataphoric, and non-directional references. It also aims to identify differences in the usage of endophoric markers in English between L2 novice writers and experienced L1 academic writers, providing insights into trends and patterns in the employment of endophoric markers in academic writing in a way that accounts for disciplinary and linguistic factors. The results reveal higher endophoric marker frequencies in the Master's Thesis Corpus, suggesting distinctive discourse patterns among Czech learners using English as an L2, with a predominance of specific endophoric markers and notable cross-disciplinary variation.

Resumen

El objetivo de este estudio es contribuir al análisis del discurso intercultural e interdisciplinar, arrojando luz sobre las prácticas metadiscursivas de los estudiantes de inglés como segunda lengua. La investigación, basada en un corpus especializado de Trabajos de Fin de Máster en lengua inglesa escritos por estudiantes universitarios checos, explora el uso de marcadores endofóricos y sus características. Para permitir la comparación interlingüística e intercultural, se compiló un corpus de referencia que representa el discurso académico en inglés de hablantes nativos.

Se utiliza una nueva taxonomía que, basada en la lista de marcadores endofóricos de Hyland (2005) y en una extensa revisión bibliográfica, clasifica los marcadores endofóricos en tres grupos: puramente direccionales, marcadores que utilizan palabras específicas y una categoría que combina los dos primeros. El estudio explora dónde cuándo y cómo aparecen estos marcadores y la función que tienen, centrándose en las referencias anafóricas, catafóricas y no direccionales. También pretende identificar las diferencias que hay en el uso de marcadores endofóricos entre escritores noveles para quienes el inglés es su segunda lengua y escritores académicos expertos para quienes el inglés es su primera lengua. Así mismo, tiene como objetivo proporcionar información sobre tendencias y patrones en el empleo de marcadores endofóricos en la escritura académica en lengua inglesa, teniendo en cuenta la influencia en su uso de factores disciplinares y lingüísticos.

Los resultados revelan una mayor frecuencia de marcadores endofóricos en el corpus de Trabajos de Fin de Máster, lo que sugiere patrones discursivos distintivos entre los estudiantes checos que utilizan el inglés como segunda lengua, con predominio de marcadores endofóricos específicos y notables variaciones entre disciplinas.

Palabras clave: metadiscurso, marcadores endofóricos, Trabajos de Fin de Máster (TFM), discurso académico, análisis intercultural.

1. Introduction

Over the past two decades, metadiscourse has received considerable attention in English for academic purposes research. Metadiscourse markers in academic genres have been analysed using Hyland's (2005) classification scheme and in the models

published by Mauranen (1993) and Ädel (2006). Endophoric markers have been examined alongside other features of metadiscourse across various genres, languages, and disciplines, as well as over time. This study adopts Hyland's (2005) interpersonal model of metadiscourse, which distinguishes between the interactive and interactional categories of metadiscourse. Interactional metadiscourse devices aim to engage the reader with the text (e.g. hedges, boosters, attitude markers), while interactive metadiscourse markers aim to guide the reader through the text (e.g. frame markers, endophoric markers, evidentials).

Endophoric markers play a crucial role as metadiscourse devices, enabling writers to reference elements within the same text (e.g. *as mentioned above*, *in Section 3 below*). This, in turn, makes it easier for readers to make connections between different parts of the text and comprehend the author's argumentative structure. As Hyland states, these markers "make additional ideational material salient and therefore available to the reader" (2005: 51). By "referring to earlier material or anticipating something yet to come", writers help the reader "steer to a preferred interpretation or reading of the discourse" (2005: 51).

It is widely acknowledged that endophoric markers are crucial metadiscourse elements frequently employed in various forms of writing, such as books, research articles (RAs), student essays, Master's theses (MTs), and PhD dissertations. They act as signposts, guiding readers to specific facts, examples, methods, and research findings presented elsewhere in the text. However, endophoric markers are typically studied in conjunction with other categories of interactive metadiscourse, rather than as a distinct metadiscourse category.

Cross-linguistic variations in the use of metadiscourse including endophoric markers have been studied in RAs, MTs and essays written in English and other languages, such as Chinese (Kim and Lim 2013; Mu et al. 2015), Spanish (Mur-Dueñas 2011; Lee and Casal 2014), Catalan and Spanish (Martín-Laguna and Alcón 2015) and French and Norwegian (Dahl 2004), among others.

Several studies have explored metadiscourse use in learner academic writing in English, comparing L1 and L2 users and various proficiency levels (e.g. Liao 2020). Ädel (2006) discovered that Swedish students tended to overuse metadiscourse markers in their L2 English essays compared to American and British university students, indicating potential issues with communicative competence. In contrast, Burneikaitė's (2008, 2009) study comparing L1 and L2 MTs in English by students in Britain and Lithuania produced different results. The underuse of endophoric markers by L2 Lithuanian students was argued to be a consequence of their "inexperience in structuring lengthy texts" (Burneikaitė 2008: 45). Kobayashi (2017) investigated the developmental patterns of metadiscourse in L2 English essays by Asian learners with diverse L1 backgrounds,

17

identifying varying frequencies of metadiscourse use across proficiency levels. These differences were attributed to the learners' L1 rhetorical strategies rather than their lexical or grammatical competence or level of proficiency.

A wide range of cross-disciplinary studies have revealed varying frequencies and uses of endophoric markers in complete RAs (Cao and Hu 2014; Hyland and Jiang 2020) as well as in RA introductions (Del Saz Rubio 2011; Kawase 2015) and abstracts (Khedri et al. 2013; El-Dakhs 2018). Cao and Hu (2014) studied the use of endophoric markers in quantitative and qualitative RAs in three soft disciplines (applied linguistics, education and psychology). Clear cross-disciplinary differences were identified: the applied linguistics and education RAs employed linear references (i.e. forward or backward references) more frequently than the psychology RAs. Nevertheless, on the whole non-linear endophoric reference dominated (e.g. *see Table 1, in Fig. 2, as demonstrated in Excerpt 3*). Hyland (2005) noted that within soft disciplines, endophoric markers serve as a mechanism to enhance discourse, with a primary goal of facilitating readers' swift and direct access to relevant information distributed across various sections of the text.

In recent years, there has been significant attention given to the evolving understanding of interactive metadiscourse in academic writing, particularly in RAs, with a focus on its use across diverse disciplines, both soft and hard. Khedri et al. (2013) analysed 60 RA abstracts in applied linguistics and economics; Dahl (2004) examined 180 RAs in medicine, economics, and linguistics; Khalili and Aslanabadi (2014) studied 20 RAs in dentistry; and Celiešienė and Sabulyte (2018) compared metadiscourse marker usage in RAs in informatics, energy, and civil engineering. The results indicated that hard science texts employ more visual elements, which accounts for the frequent use of endophoric markers and code glosses. This practice ensures effective signposting and accurate interpretation by the reader of the information provided. Another cross-disciplinary research study compared Master's and doctoral dissertations across six disciplines written in English by L2 Hong Kong students (Hyland 2004). The results revealed that students in the hard disciplines, particularly those in engineering, demonstrated a relatively higher use of endophoric markers. This highlighted "the multi-modal nature of scientific discourse" (Hyland 2004: 147) and the increased dependence on arguments requiring frequent reference to elements such as tables, figures, examples, and photographs. Hyland and Jiang (2020) conducted a diachronic study analysing a corpus of 2.2 million words from articles in top journals across four academic disciplines, aiming to explore changes in interactive metadiscourse usage since 1965. The results indicate an increased usage of endophoric markers in English across diverse disciplines over the past 30 years. This upward trend seems to reflect a growing tendency for texts to become more explicit, employing

techniques like exemplification, illustration, and referencing tables and graphical data to clarify associations (Hyland and Jiang 2020).

Research into endophoric markers extends beyond RAs and includes other genres such as undergraduate essays (Ädel 2006), MTs (Lee and Casal 2014), postgraduate dissertations (Hyland and Tse 2004; Kawase 2015), and university textbooks (Hyland 1999). Bunton's (1999) study on postgraduate dissertations revealed that higher level endophoric references, i.e. those referring to larger portions of text or operating over greater distances, were more frequent. This phenomenon was attributed to the length of the texts. However, considerable variations among writers were also reported.

To date, research on Czech English-medium learner discourse has been limited to a handful of studies that concentrate on individual metadiscourse markers, such as sentence linkers (Vogel 2008; Povolná 2012), code glosses (Guziurová 2022), attitude markers (Jančaříková 2023) and self-mention (Dontcheva-Navrátilová 2023). However, there have been no investigations into the use of endophoric markers by Czech students in English.

As noted above, few studies to date have focused on metadiscourse in L2 MTs across soft disciplines. This paper aims to fill this gap by examining endophoric markers in L2 MTs written in English by Czech university students. Two corpora were compiled for this purpose: the Master's Thesis Corpus (MT_LLE) and the Research Article Corpus (RA_LLE). These corpora were used to compare and contrast the use of endophoric markers by English L2 and L1 writers, examining differences across levels of writing expertise and linguacultural dimensions. Additionally, the realisation of endophoric markers is compared across three disciplines, i.e. linguistics, literary studies, and education. These disciplines are integral to English studies at Czech universities, and the MTs included in the L2 corpus focus on these specific areas. A new taxonomy was devised and used in the analysis (for more details concerning the taxonomy, see Section 2.4).

This study seeks to answer the following research questions:

1. What is the frequency of occurrence, realisation, and function of endophoric markers in English-medium L2 (Czech) MTs and L1 published RAs?

2. What are the similarities and differences in the frequency of occurrence, realisation, and function of endophoric markers in English-medium L2 (Czech) MTs and L1 published RAs?

3. What are the similarities and differences in the frequency of occurrence, realisation, and function of endophoric markers in the three disciplines in English-medium L2 (Czech) MTs and L1 (English) published RAs?

19

2. Data and Methodology

2.1. Corpus Characteristics and Compilation

For the present study, two distinct corpora were compiled: the Master's Thesis Corpus (MT_LLE) and the Research Article Corpus (RA_LLE). Each of these corpora will be discussed in the following section, accompanied by an explanation of the selection criteria applied.

The analysis of endophoric markers was conducted on a specialised learner corpus of English-medium MTs written by Czech university students majoring in English language and literature, either at the Faculty of Arts or the Faculty of Education of Masaryk University in Brno. The corpus comprises theses written and submitted from 2010 to 2018. The following additional criteria were applied: the authors' native/first language is Czech (L1), and only theses awarded high grades, that is, 'A' ('Excellent'/'Merit'), were included. A total of 48 MTs were included in the corpus, distributed equally across three academic disciplines: linguistics, literary studies, and education, with 16 theses drawn from each discipline. Prior to the analysis, all MTs underwent a refinement process, which involved the removal of elements such as abstracts, citations, tables, and figures. This approach ensured that the analysis focused exclusively on the students' written discourse in running text found within the MT macrostructure. The corpus comprises a total of 948,000 words.

In order to investigate typical elements of metadiscourse in the MT corpus, it was necessary to compile a reference corpus representing L1 English academic discourse. Due to the fact that there is no corpus of English L1 MTs available, the reference corpus (RA_LLE) comprises RAs authored by scholars who are native English speakers and affiliated with universities in Britain or the United States. The authors' native speaker status (L1) is indicated by their names, affiliations, and CVs. The selection process exclusively considered articles from the same three fields or disciplines as in the MT corpus, i.e. linguistics, literary studies, and education. These articles were single-authored, produced between 2010 and 2018, and published in widely recognised international journals indexed in the Web of Science database. Linguistics articles were sourced from *Discourse & Communication*, the *Journal of Pragmatics*, and *Applied Linguistics*. Literary studies articles were selected from *Eighteenth-Century Fiction*, *New Literary History*, and *SEL: Studies in English Literature*. Articles focusing on educational themes were taken from *Language and Education*, *Language Teaching Research*, and *Language Learning*. The corpus comprises 36 articles (12 articles per discipline), all of which have undergone the same cleaning process applied in the MT_LLE corpus. The total word count for this corpus is 243,000 words (see Table 1 below).

Corpus	No. of texts	Word count	Disciplines
MT_LLE	48	948,000	Linguistics, literary studies, education
RA_LLE	36	243,000	Linguistics, literary studies, education

Table 1. Composition of the MT_LLE and RA_LLE

2.2. Corpus Comparison and 'Tertium Comparationis'

Due to the unavailability of MTs authored by L1 English writers, no reference corpus could be compiled. Flowerdew (2015) aligns with Tribble (2002) in suggesting that in the absence of a directly comparable corpus, concessions must be made, and an 'analogue' corpus can be used. This refers to a corpus that closely resembles student writing in terms of genre and discipline (Flowerdew 2015: 60). The approach adheres to the 'tertium comparationis' criteria advocated by Connor and Moreno (2005: 155; see also Moreno 2008: 25). Therefore, the reference corpus of RAs (RA_LLA) was carefully compiled to mirror the MT_LLE corpus as closely as possible. Despite the differences between MTs and RAs in their aims, scope, audience, and criteria (Paltridge 2002; Flowerdew 2015), it has been argued that "there are significant overlaps in lexico-grammar and rhetorical functions" (Flowerdew 2015: 60). As certain sections of MTs (or dissertations) may eventually evolve into RAs (Swales 1990: 178), previous research on metadiscourse has often juxtaposed MTs with RAs, serving as potential templates for inexperienced writers (see, for instance, Koutsantoni 2006; Pujol Dahme and Selfa Sastre 2015; Abdollahzadeh 2019; Qiu and Ma 2019).

The MT_LLE corpus and RA_LLE corpus are fully comparable in terms of disciplines (linguistics, literary studies, and education), as well as the time span in which they were written (2010-2018). The analysis of the corpora thus incorporates the concept of intercultural, cross-linguistic and cross-disciplinary variation.

While it is worth noting that the sizes of these corpora vary, the primary focus of this study is to explore the occurrence of endophoric markers in theses written in English by Czech native speakers. To mitigate the discrepancies in corpus size, the research findings were normalised to occurrences per 10,000 words, and the log-likelihood ratio test was applied to account for size differences. This approach allows for drawing reliable conclusions while acknowledging the inherent limitations of the differing corpus sizes.

2.3. Procedure

Both corpora were compiled and searched using SketchEngine software (Kilgarriff et al. 2004). Statistical significance was determined using the non-parametric log-

likelihood test, following common practice in specialised corpus analysis. A significance level of <0.05 was established, with very low p-values indicated as <0.001.

Based on Hyland's (2005) list of endophoric markers and a comprehensive review of the relevant literature, a list of endophoric markers was compiled and searched within the corpora. After this initial stage, additional items identified as potential endophoric markers (e.g. *extract*, *excerpt*, *here*) were added to the list. It was required that every instance of potential endophoric markers be examined in context to ensure that they functioned as expected. According to Hyland and Tse (2004), who drew on Bunton's (1999) distinction between 'research acts' and 'writer acts', each potential discourse marker must be assessed to determine whether it carries propositional value (Example 1) or metadiscoursal value (Example 2) (Hyland and Tse 2004: 166).

(1) It was not enough for the teachers to value students' voices in the discussion; many of them felt they had to vanish completely, *following* procedures for discussion that completely removed the teacher from participation. (RA_EDU_04)

(2) *The following extracts* show other instances of Cherry's praise of Canadian-born ice hockey players while at the same time reiterating their Canadian regional allegiance and affinity. (MT_LIT_03)

The token *following* is used as a metadiscoursal device in Example 2; however, in the context of Example 1, it signals discourse-external relations.

Furthermore, within the dimension of interactive metadiscourse, careful attention has to be given to certain markers, which could be assessed either as frame markers or as endophoric markers, based on the context. Below are two examples of the token *analysis* in context (Examples 3 and 4):

(3) I turn now to an *analysis* of what these contrasts mean in terms of language use. (RA_LIN_12)

(4) Nevertheless, as demonstrated *in the analysis*, considerable attention to the choice of lexico-grammatical items needs to be paid. (MT_LIN_09)

In Example 3, the word *analysis* introduces or frames the content that follows and is typically categorised as a frame marker, labelling the stage of the discourse based on Hyland's theory of metadiscourse. In Example 4, the phrase *in the analysis* functions as an endophoric marker, referring to the specific analysis within the same text for content and clarification.

2.4. Taxonomy

When analysing endophoric markers, authors typically provide a list of such markers or offer a few examples (Hyland and Tse 2004; Hyland 2005). Bunton (1999), influenced by Mauranen's (1993) research on metatext and text reflexivity and by Crismore and Farnsworth's (1990) terms of reviews and previews, created a taxonomy of endophoric markers. He referred to them as 'text references' and proposed a taxonomy that distinguishes between linear and non-linear text references. Linear text references are explicit references to other parts of the linear text. They are further categorised based on the direction of the reference as reviews "looking back, repeating, summarising or referring to an earlier stage of the text" (e.g. *as noted earlier in the paper*), previews "looking forward, anticipating, summarising or referring to a later stage of the text" (e.g. *the next section*), and overviews "looking in both directions, referring to the current stage of the text in overall terms" (Bunton 1999: 45). It is worth noting that 'overviews' in this context concern references that consider the text from a broader perspective —e.g. "This was only used by X among the (subjects) examined in this thesis" (Bunton 1999: 46). Bunton also showed the level of reference (scope), which may refer to the written discourse as a whole or to a specific chapter. Additionally, he explored the distance to the text segment being referred to. Non-linear text references are explicit references to *tables*, *figures*, *charts*, or *appendices*.

This taxonomy, expanded using Hyland's (1999, 2005) theory of metadiscourse to distinguish between endophoric and frame markers, served as a valuable framework in Cao and Hu's (2014) analysis of RAs. Burneikaitė (2009) incorporated Bunton's (1999) classification of metatext into her work, which resulted in a classification that divides endophoric markers into non-linear and linear text references. As with Bunton's taxonomy, non-linear text references cover elements like *appendices*, *tables*, *figures*, *charts*, *graphs*, and *diagrams*. Concerning linear text references, the author classified them into several levels: thesis-level markers (e.g. *paper, study, thesis, analysis, dissertation, research (work), essay, article, project*), chapter/section-level markers (e.g. *chapter, (sub)section, part, paragraph*), sentence-level markers (e.g. *example, instance, illustration, sentence, case*) and vague markers (e.g. *above, below, further, following, previously, what follows, remainder*) (Burneikaitė 2009: 13).

In an attempt to capture and categorise all instances of endophoric markers identified in the corpora. I propose the following taxonomy. While it has some similarities with Bunton's taxonomy, it is primarily based on Hyland's definition of metadiscourse and the interactive dimension, taking into account additional elements identified in the data.

23

I partially agree with the decision to divide endophoric markers into linear and non-linear elements. However, when analysing the corpora, I came across instances of endophoric markers that might be classified as non-linear according to existing taxonomies, although their context and function indicated otherwise. For example, the marker *table* in the following sentences does not qualify as an example of non-linear reference since it refers back or ahead (see Examples in 5).

(5) As we have seen, the focus on significance also leads researchers astray in suggesting that nonsignificant findings should not even be reported, *as in Table 1 above*; not reporting full findings only helps to further obscure what was observed in the data and why patterns might or might not have been statistically significant in the first place. (RA_EDU_11)

Table 1 below represents only a sample of relevant verbs (i.e. verbs related to the issue of existence or appearance on the scene) as presented in the book. (MT_LIN_02)

For that reason, I abandoned the distinction between linear and non-linear and in its place propose the following taxonomy (see Table 2).

Endophoric markers are categorised into three groups: purely directional markers, markers using specific words, and a category that combines the first two. Markers using specific words can refer to the entire discourse (*article, essay, here* [non-directional reference], *paper, study, thesis, work*), to individual parts of the discourse (*analysis, chapter/subchapter, conclusion, discussion, introduction, page, paragraph, part, section/subsection*) or to items incorporated into the discourse (*appendix, chart, diagram, example, excerpt, extract, figure, here* [either as backward or forward reference], *table*). These markers can refer back or ahead by the use of tenses (Examples in 6) or immediate context (Example 7).

(6) *In the examples analysed so far, I have shown* how the interview is designed to foreground the IE's feelings and thoughts, to be collaborative rather than confrontational, and to establish a version of events in terms of what happened to her personally. (RA_LIN_01)

These key concepts *will be applied in the analysis* of online weight loss advertising. (MT_LIN_13)

(7) The items in the FLCAS (Horwitz et al., 1986) reflect the above mentioned characteristics of an anxious student, as we can see *in these examples*: <EXAMPLE>, or <EXAMPLE>. (MT_EDU_09)

Moreover, they can also serve as non-directional endophoric markers, pointing to the current discourse, section, or incorporated item itself (Examples in 8). When combined with clearly directional markers, thus labelled as combined endophoric markers, they can, of course, serve as anaphoric and cataphoric signposting markers (Example 9).

Categories of EM	Subcategory	Type of reference	Examples of EM	Examples of EM in sentences from the corpora
Purely directional		Anaphoric	*above, aforementioned, before, earlier, previously, just, already*	For example, *as mentioned above*, many of the participants in this study viewed discussion both as a community-building exercise as well as an assessment. (RA_EDU_04)
		Cataphoric	*below, following/ as follows, further on, next, later (on), subsequently*	*Below* are some of the most frequent and notable references appearing in speeches of American presidents. (MT_LIN_01)
Specific	Reference to the entire discourse	Anaphoric	*article, essay, here (*non-directional reference*), paper, study, thesis, work*	My hope is that the issues addressed *in this paper* have at least raised some awareness regarding statistical significance testing, its accurate conduct, and its limitations, and that L2 quantitative research will benefit to some small, though perhaps significant, degree as a result. (RA_EDU_11)
		Cataphoric		As *this article* will reveal, in the early nineteenth century laughter is elevated to the condition of the language of feeling and indicative of Romantic genius. (RA_LIT_10)
		Non-directional		*This thesis* works with the concept of new racism, as dealt with in CDA by Dijk (e.g. 1991, 2000, 2002) or Wodak (2001) or by Romm (2010) and Kundnani (2007) in social sciences. (MT_LIN_10)
	Reference to individual parts of the discourse	Anaphoric	*analysis, chapter/ subchapter, conclusion, discussion, introduction, page, paragraph, part, section/subsection*	As mentioned in *Chapter 6*, the type of image that appears in the discourse is very significant and constitutes a part of the information load of the advertisement. (MT_LIN_13)
		Cataphoric		The deployment of such expressions thus appears to be doing other kinds of interactional work, a point which we will return to discuss in further detail *in section four*. (RA_LIN_06)
		Non-directional		It is far beyond the scope of *this section* to provide a comprehensive review of research on motivation and subsequent theories explaining the results. (MT_EDU_10)

25

Categories of EM	Subcategory	Type of reference	Examples of EM	Examples of EM in sentences from the corpora
Specific	Reference to items inserted into the discourse	Anaphoric	*appendix, chart, diagram, example, excerpt, extract, figure, here* (either as backward or forward reference), *table*	Consider the following metaphor: <EXAMPLE> *In this excerpt*, the speaker suggests that war has human-like qualities, demanding people change their attitude and act even when they do not want to. (MT_LIN_01)
		Cataphoric		Verbs that have been detected as the only representatives of a certain class will be included *in Table 15*. (MT_LIN_02)
		Non-directional		Alexander Scourby was an obvious choice to narrate the epic project for the American Foundation for the Blind (*see Figure 1*). (RA_LIT_08)
Combined (directional + specific)		Anaphoric	*e.g. above, preceding, previous + article, section, figure*	As with *previous excerpts*, Peng's positioning as a potential target of discrimination emerges in the spaces created by the interviewers' questions, and thus the ambiguity of his representational positioning results from its sequential location in the co-constructed production of talk. (RA_LIN_09)
		Cataphoric	*e.g. below, following, next + thesis, part, table*	More details on this issue are provided *in the following sections*. (MT_EDU_03)

Table 2. Taxonomy of endophoric markers with corpus examples

(8) However, the author is aware of certain weaknesses in her lesson plans that are defined *in the discussion*. (MT_EDU_16)

In this introductory section, I start by clarifying the distinction between first and second order concepts of im/politeness, briefly trace the history of mock politeness within impoliteness studies and present the definition of mock politeness which will be employed in this paper. (RA_LIN_05)

While on the surface some of these items may not seem inherently tied to a rhetoric of uncertainty and doubt, closer inspection of concordance lines and texts (*see Table 3*) reveals the rhetorical function each serves in the discourse. (RA_LIN_04)

(9) As *the previous discussion* has implied, curricular thinking has been more the exception than a well-honed practice in ISLA. (RA_EDU_08)

> This is even more evident *in the following subchapters (6.2 and 6.3)* where particular models of instructional design are introduced. (MT_EDU_14)

Directional markers include anaphoric (e.g. *above, aforementioned, before, earlier, previously, just, already*) and cataphoric (e.g. *below, following/as follows, further on, next, later (on), subsequently*) markers. These can be used, for example, in adverbial phrases (thus labelled as purely directional markers) (Example 10) or in nominal phrases, usually in combination with specific markers (but not exclusively, as in Example 11), thus labelled as combined markers, where they function as premodifiers (Example 12) or postmodifiers (Example 13).

(10) *As mentioned earlier*, the topic of weight loss can be considered as a very sensitive issue because it is closely connected with the concept of ideal body image and lifestylism. (MT_LIN_13)

And *later on*, this distinction will lead me to navigate between ways of reading for description and the varied ambitions of so-called "descriptive reading" as such. (RA_LIT_05)

(11) Together with *the above stated definition*, this settles the notion of error clear enough for the purposes of this work. (MT_EDU_07)

However, the design of her extended response from line 8 onwards quickly shifts to a more experiential display of emotionality, including the same *features previously noted in section 3 above*. (RA_LIN_01)

(12) We have described the role of experiments, measures, corpora, and basic skills tutors *in previous sections*. (RA_EDU_12)

(13) As can be derived from *the example above*, it is not sufficient to teach the students only the separate concepts of language. (MT_EDU_02)

3. Results and Discussion

The results of the quantitative analysis of endophoric markers in the two corpora are summarised in Table 3. The frequency of endophoric markers found in the learner corpus, consisting of MTs written in English by Czech university students, was significantly higher (41.6 per 10,000 words) than in the RA_LLE reference corpora, with 28.7 occurrences per 10,000 words. This difference was found to be statistically significant (LL test, p-value <0.001).

Endophoric marker categories	Subcategories	MT_LLE		RA_LLE		LL G2	p-value
		n	**pttw**	**n**	**pttw**		
Purely directional		608	6.4	60	2.5	64.6176	<0.001
Specific		2,079	21.9	509	21.0	0.8714	0.3506
	Reference to the entire written discourse	234	2.4	132	5.4	47.8588	<0.001
	Reference to individual parts of the written discourse	927	9.8	69	2.9	140.9983	<0.001
	Reference to items incorporated into the written discourse	918	9.6	308	12.7	15.9739	<0.001
Combined		1,256	13.2	128	5.3	127.0166	<0.001
TOTAL		**3,943**	**41.6**	**697**	**28.7**	**89.3992**	**<0.001**

Table 3. General frequency of endophoric markers across the corpora (Abbreviations: n: number; pttw: per ten thousand words)

Comparing the three categories of endophoric markers (purely directional, specific, and combined), the highest frequency of endophoric references in both corpora was associated with specific endophoric markers. The frequency of specific markers was nearly the same in the MT_LLE and RA_LLE corpora (21.9 and 21.0, respectively; LL test, p-value 0.3506), which may indicate a genre-related trend in research writing. This aspect will be examined further in a cross-disciplinary analysis of specific endophoric references later in this section.

Another salient observation regarding specific reference and its three subcategories is the frequency of referencing items integrated into the discourse. This referencing frequency was notably higher in the RA_LLE corpus (12.7 occurrences) compared to the MT_LLE corpus (9.6 occurrences), and this difference was found to be statistically significant (p-value < 0.001). In contrast, combined endophoric markers were more prevalent in the MT_LLE corpus (13.2) than in the RA_LLE corpus (5.3), and again this difference was statistically significant (p-value < 0.001). Similar results have been found in engineering RAs, where a greater density of lexical bundles occurred (e.g. *is shown in Figure*) (Hyland 2008). This may suggest that Czech undergraduate students tend to be exceptionally clear or meticulous when directing their readers to specific parts of their discourse in

English-medium theses. They emphasise this by using multiple endophoric markers simultaneously, as illustrated in Example 14.

(14) *It has been mentioned in the opening paragraph of this section (see page 49)* that in his novel Ellis attributes the rampant indifference that leads to the collapse of communication not only to the influence of television but also of glossy magazines. (MT_LIT_05)

Within each category of endophoric markers (i.e. purely directional, specific, and combined), endophoric markers can be categorised as anaphoric or cataphoric. Additionally, specific reference markers can also be classified as non-directional (see Table 2).

Concerning the overall distinctions among anaphoric, cataphoric, and non-directional references, Table 4 provides a summary of the data in both the RA_LLE and MT_LLE corpora. Anaphoric reference predominates over cataphoric in both corpora, as authors tend to revisit previously mentioned content, reinforcing the reader's recollection of already presented knowledge. Anaphoric reference occurs most frequently among Czech L2 English student writers, at a rate of 16.3 occurrences per 10,000 words, while L1 English expert writers predominantly use non-directional reference, with a frequency of 13.2 instances per 10,000 words. This difference is statistically significant (p-value < 0.001). However, the difference in the use of non-directional reference between the MT_LLE and RA_LLE corpora was not found to be statistically significant (p-value 0.0952).

	MT_LLE n	MT_LLE pttw	RA_LLE n	RA_LLE pttw	LL G2	p-value
Anaphoric reference	1,548	16.3	246	10.7	54.4399	<0.001
Cataphoric reference	1,273	13.4	131	5.4	126.7413	<0.001
Non-directional reference	1,122	11.8	320	13.2	2.7844	0.0952

Table 4. Anaphoric, cataphoric and non-directional endophoric markers across the corpora

In Lee and Casal's (2014) investigation of engineering MTs, the results indicated that L1 English writers directed readers' attention to previously and subsequently discussed material equally. In contrast, Spanish L1 authors primarily included vague anaphoric reference markers. It is worth noting that Czech students use all three types of reference at comparable frequencies, ranging from 11.8 to 16.3 occurrences in their English-medium theses. Expert English L1 writers, conversely,

exhibit a clear preference for non-directional and anaphoric reference, with frequencies of 13.2 and 10.7 occurrences per 10,000 words, respectively.

Anaphoric markers, which were used both purely directionally and in combined phrases, were counted collectively, and the comparison of their frequency across the corpora can be seen in Figure 1. The most frequently used anaphoric markers in all three corpora were the words *above* (frequency of 4.8 in MT_LLE and 2.0 in RA_LLE) and *previous/ly* (2.0 in MT_LLE and 0.7 in RA_LLE). This aligns with the findings of Hyland and Jiang (2020), who identified *above* as the most frequently used linear marker in both 1965 and 1985. The third most frequently used anaphoric marker in the MT_LLE corpus was *already*, occurring at a rate of 1.4, in contrast to its significantly lower incidence in the RA_LLE corpus, where its frequency was 0.2. This vague anaphoric reference is not commonly used in English; however, in Czech, the formal phrase "jak *již* bylo zmíněno" [as has *already* been mentioned] is a frequently used anaphoric marker within academic discourse. Therefore, the more frequent use of the marker *already* by Czech students when writing in English might be attributed to cross-linguistic influence from the Czech language.
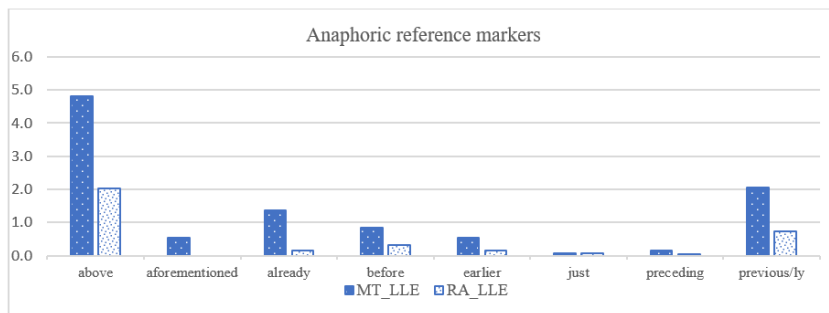
Figure 1. Anaphoric reference markers (purely directional and combined, counted together)

Figure 2 illustrates the distribution of cataphoric reference markers across the corpora. The two most commonly used cataphoric markers were consistent across both corpora: *following* (frequency of 6.0 in MT_LLE and 2.2 in RA_LLE) and *below* (1.8 in MT_LLE and 1.3 in RA_LLE). Czech students and expert writers used *next* at a similar frequency (0.6 and 0.5, respectively). Czech students also employed the marker *further (on)*, but this marker did not appear in the RA_LLE corpus. A similar pattern was observed with the marker *subsequent/ly*, which was absent from the RA_LLE corpus.
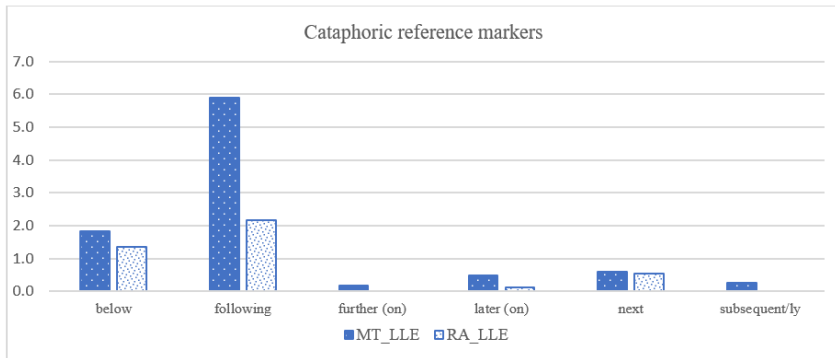
Figure 2. Cataphoric reference markers (purely directional and combined, counted together)

The following part of the results section will examine cross-disciplinary variations regarding the occurrence, realisation, and frequency of endophoric markers. A more in-depth exploration is provided for specific endophoric markers and their three subcategories, as these markers were the most frequent form of endophoric reference across both corpora.

Table 5 presents the distribution of references to the entire discourse, drawing data from the category of endophoric reference using specific expressions. When referencing their written discourse, authors predominantly employed terms such as *thesis* and *work* in MTs, and *article* and *study* in RAs. The second most frequently used term in the MT_LLE was *here*, which ranked third in the RA_LLE corpus. However, it had a higher frequency than in MT_LLE, with 1.2 occurrences per 10,000 words. *Here* functioned as a self-reference word to denote the thesis or article itself. In parts of the corpora it was challenging to discern whether the reference pointed to the entire discourse or a specific part (e.g. a section) (see Examples 15 and 16). Consequently, all instances of *here* referring to the entire discourse or its parts were included in the count.

(15) The interview excerpts considered *here* come from a larger ethnographic project involving individuals who immigrated to a small Midwestern city in the US, which I have named Midtown, and who were enrolled in free ESL classes offered by a social services agency. (RA_LIN_09)

(16) In conclusion I would like to say that I hope I have correctly grasped the issues compiled in this section, however, I rather apologize *here* for any prospective misinterpretation. (MT_LIT_04)

Reference to the entire written discourse	MT_LLE		RA_LLE	
	n	pttw	n	pttw
article	0	0.0	42	1.7
essay	0	0.0	8	0.3
here	41	0.4	29	1.2
paper	20	0.2	18	0.7
study	19	0.2	32	1.3
thesis	119	1.3	0	0.0
work	35	0.4	3	0.1
TOTAL	**234**	**2.5**	**132**	**5.4**

Table 5. Distribution of reference to the entire written discourse (in specific category only)

As mentioned above, endophoric reference to individual parts of the discourse and to items incorporated into the discourse reveals similarities in the research-oriented genre (theses and articles). Therefore, a closer examination of these two subcategories, especially investigation of cross-disciplinary variations among the three disciplines (linguistics, literary studies, and education), would be insightful.

Table 6 summarises the distribution of references to individual parts of the written discourse, drawing data from the specific-only and combined categories and counting them together. The relative frequencies clearly indicate that referencing individual parts of the written discourse was significantly higher in MT_LLE than in RA_LLE across all three disciplines, possibly explained by the length of the research genre. While references to *chapters* and *subchapters* were the most frequent endophoric markers across all three disciplines of MT_LLE, such references were naturally absent in RA_LLE, as these are not part of the texts. Reference to *sections* and *subsections* occurred across all three disciplines, with the incidence consistently higher in MT_LLE than in RA_LLE.

Concerning frequency of references to specific parts of discourse across disciplines, linguistics showed the highest occurrence of endophoric reference in both MA_LLE and RA_LLE (19.7, 9.4 respectively). However, upon closer examination of specific references to *analysis* and *discussion*, which appear in both research genres and are considered core parts of each research study, clear differences emerge.

Reference to *analysis* was more frequent in RAs than in MTs in linguistics and education studies, highlighting the significance of analysis sections in RAs, as recognised by expert writers. Reference to *discussion* was also more frequent in RAs in linguistics, with the same frequency found in MTs and RAs in education.

Reference to *pages*, frequently mentioned in lists of endophoric markers in various studies of RAs (cf. Hyland 2005; Lee and Casal 2014), was absent from RAs in all three disciplines, and in the MT corpus only 19 instances were found.

Based on the description so far and the contents of Table 6, it is evident that reference to individual parts of the discourse was lowest in literary studies MTs and RAs. This is explained by the focus of the discipline itself; as analysis and discussion are distributed throughout the discourse in literary studies, there is no clear tendency to refer to them explicitly.

Reference to individual parts of the written discourse	Linguistics		Literary studies		Education	
	MT_LIN	RA_LIN	MT_LIT	RA_LIT	MT_EDU	RA_EDU
	pttw	pttw	pttw	pttw	pttw	pttw
analysis	3.3	4.0	1.1	0.0	0.2	0.4
chapter/subchapter	4.4	0.0	4.1	0.0	8.8	0.0
conclusion	0.3	0.1	0.1	0.0	0.1	0.0
discussion	0.6	0.8	0.4	0.1	0.2	0.2
introduction	0.1	0.1	0.2	0.0	0.1	0.0
page	0.2	0.0	0.5	0.0	0.0	0.0
paragraph	0.3	0.0	0.4	0.0	0.5	0.0
part	4.4	0.1	1.3	0.4	3.0	0.1
section/subsection	6.1	4.2	0.8	0.6	2.2	1.2
TOTAL	**19.7**	**9.4**	**8.8**	**1.1**	**15.0**	**1.9**

Table 6. Distribution of reference to individual parts of the written discourse (data taken from specific and combined category, counted together)

33

Table 7 illustrates the distribution of references to items incorporated into the discourse, with data taken from the specific-only and combined categories and counted together. Once again, endophoric markers are used significantly less in the discipline of literary studies in both MTs and RAs. The highest frequency was identified in linguistics studies in both MTs and RAs (with the frequency almost equal), while in the discipline of education, RAs displayed much higher use of these markers than MTs (15.2 to 9.6, respectively).

Reference to items incorporated into the discourse	Linguistics		Literary studies		Education	
	MT_LIN	**RA_LIN**	**MT_LIT**	**RA_LIT**	**MT_EDU**	**RA_EDU**
	pttw	**pttw**	**pttw**	**pttw**	**pttw**	**pttw**
appendix	1.6	0.3	0.1	0.0	2.2	0.2
chart	1.5	0.0	0.0	0.0	0.8	0.0
diagram	0.2	0.0	0.0	0.0	0.1	0.0
example	8.9	5.8	0.6	0.4	1.0	3.4
excerpt	0.5	4.9	0.6	0.0	0.0	1.9
extract	2.5	7.9	0.1	0.0	0.0	0.0
figure	2.3	4.8	0.0	0.4	2.6	2.0
here	3.0	1.0	0.3	0.4	0.4	1.6
table	7.8	1.9	0.0	0.0	2.6	6.0
TOTAL	**28.2**	**26.8**	**1.7**	**1.1**	**9.6**	**15.2**

Table 7. Distribution of reference to items incorporated into the discourse (data taken from specific and combined category, counted together)

Regarding the actual words used in the endophoric markers, the most frequently used references to incorporated items in linguistics MTs were *example*, *table*, and *here*, whereas in linguistics RAs, they were *extract*, *example*, *excerpt*, and *figure*. In literary studies MTs and RAs, the most frequent references were *example* and *here*. The endophoric marker *here* is used either as an anaphoric marker (Example 17) or a cataphoric marker (Example 18), often substituting for words such as *example*, *extract*, *excerpt*, etc.

(17) Initially putting laughter and tears together, Hazlitt soon adopts the conventional approach of having tears signify pathos and laughter embodying derision: <EXAMPLE>. While laughter is not a sign of pity *here*, it is indicative of the sympathetic imagination since laughter reflects our capacity to discern what is "unreasonable and unnecessary" in others and act accordingly. (RA_LIT_10)

(18) Let me demonstrate it on a few examples *here*: <EXAMPLE>. Simple structured compliments such as <EXAMPLE> or <EXAMPLE> are often exploited by the characters as well (10.2%). (MT_LIN_11)

In the education MT corpus, the most frequently used reference words in this subcategory were *table*, *figure*, and *appendix/appendices*, while in the education RA corpus, they were *table*, *example*, and *figure*. In education RAs, references to *tables* were twice as frequent as in MTs (6.0 and 2.6, respectively). References to *figures* were almost the same in MTs as in RAs. Another distinction was that

authors of MTs across all three disciplines made more references to *appendix/ appendices*. Once again, the variation is attributed to the structural norms of RAs in different journals, where appendices are not common, while in MTs, use of appendices is frequent.

The occurrence of endophoric reference to items incorporated into discourse, especially to visual aids like *figures* and *tables*, is notably scarcer in literary studies than in linguistics and education. This scarcity can be explained by the primary focus in literary studies on descriptive analysis, introduction of and commentary on quotes from literary works or secondary sources, and presentation of important examples.

When combining references to *examples*, *excerpts*, *extracts*, and *here* in linguistics, authors in both MTs and RAs relied on them more than authors writing within the field of education (14.9 in MTs and 19.6 in RAs in linguistics compared to 1.4 in MTs and 6.9 in RAs in education). In contrast, references to findings summarised using visual materials (e.g. *charts*, *diagrams*, *figures*, and *tables*), the differences were less pronounced. Authors of RAs in education showed a higher reliance on these visual materials (8.1 in RA_EDU, 6.1 in MT_EDU), whereas in MTs this type of reference was more frequent in linguistics (11.8 in MT_LIN, 6.7 in RA_LIN).

4. Conclusion

The extent to which endophoric markers are explicitly taught in academic English writing courses varies. In many courses, the primary focus is on teaching students to structure their writing, develop arguments, use evidence, and cite sources properly. However, depending on the course's curriculum, the instructor's approach, and the students' proficiency levels, endophoric markers can be covered to some extent. These markers are crucial for ensuring coherence and cohesion in academic writing. Endophoric signposting helps readers understand the relationships between different parts of the text and navigate the content smoothly, making it particularly important in longer and more complex pieces of writing, such as the MT.

The analysis of endophoric markers across the Master's Thesis Corpus (MT_LLE) and the reference corpus, the Research Article Corpus (RA_LLE), has revealed distinctive patterns in how Czech university students majoring in English language and literature use these markers compared to L1 English-speaking writers.

A new taxonomy of endophoric markers was developed for this study. The taxonomy categorises endophoric markers into purely directional, specific, and combined markers, enabling a more nuanced understanding of how these markers

function in different contexts. The overall findings revealed that the MT_LLE corpus contained a significantly higher frequency of endophoric markers (41.6 per 10,000 words) as compared with the reference corpus (28.7 per 10,000 words in RA_LLE). This difference can be attributed to the respective length of the two types of written discourse, as described in other research studies (cf. Bunton 1999). However, discourse length cannot be the sole factor influencing the use of endophoric markers. Crismore et al. (1993) suggest that the use of metadiscourse is closely linked to socio-cultural contexts. This is consistent with the findings reported in a study by Afzaal et al. (2021) on the use of metadiscourse markers, comparing introductions of MTs written in Chinese and US universities. The lower use of metadiscourse markers found in Chinese introductions was attributed not only to the shorter length of the texts but also to socio-cultural factors. In Chinese writing, readers are required to take on a larger role in determining the relationship of specific sections within the text as a whole (Afzaal et al. 2021: 12).

The more frequent use of certain types and realisations of endophoric markers in RA_LLE suggests a clear link to the function of endophoric markers as used by expert writers. Notably, combined endophoric markers integrating specific and purely directional markers were more prevalent in the MT_LLE corpus, indicating a comprehensive approach by Czech students in guiding readers within their English written discourse, albeit suggestive of their inexperience. Conversely, specific markers referring to the entire discourse, individual parts, or items incorporated into the discourse were equally prevalent in both the MT_LLE and RA_LLA corpora (21.9 and 21.0, respectively; LL test, p-value 0.3506), suggesting a shared emphasis on signposting and guiding readers through research-genre texts. Furthermore, within this category, referencing to the discourse itself or to items incorporated into the written discourse prevailed in RA_LLE compared to MT_LLE, and the results were found to be statistically significant.

The examination of anaphoric, cataphoric, and non-directional reference highlights further differences between the corpora that cannot be explained solely by the length of the texts that make up the corpora but rather by the function of particular endophoric markers. Anaphoric reference dominates in Czech L2 English student writing, whereas English-speaking expert writers primarily employ non-directional reference, indicating a difference in the use of markers between Czech students and experienced L1 English-speaking writers. Similarly, Cao and Hu (2014) identified a prevalence of non-directional reference in their study. Therefore, academic English courses that emphasise the significance of non-directional reference, as observed in expert writing, can enhance students' proficiency in the use of English academic discourse.

Cross-disciplinary analysis also revealed variations in the use of endophoric markers. For instance, linguistics exhibited a higher occurrence of endophoric reference in both the MT_LLE and RA_LLE corpora. On the other hand, MTs in education displayed a significantly lower incidence of these markers than RAs in the same discipline. This is something that can be addressed by instructors of academic English by encouraging students to use more endophoric reference when presenting their results in the discipline of education. They could reference items integrated into their discourse, such as *tables* and *examples*, based on and closely related to either the quantitative or qualitative methods of their research study.

Academic writing instruction in English and coursebooks typically cover metadiscourse and include exercises for practising specific structures. However, endophoric markers are often treated as implicit knowledge that everyone is expected to possess and use. Consequently, there tends to be a lack of exercises with which to practice this aspect.

These findings highlight the potential for enhancing English academic writing courses, enabling students to develop skills for effectively communicating their ideas and arguments in written form. Additionally, students should be familiarised with the conventions of English academic discourse in their discipline, including the use of endophoric markers. One approach could involve students identifying endophoric markers in samples of specific text genres, thereby making their implicit knowledge about the use of endophoric markers explicit. Moreover, instruction on metadiscourse, including endophoric markers, can be integrated with introductory lessons on using corpus tools and building their own corpora. Awareness-raising activities could be employed to compare the use of endophoric markers in expert writing within their discipline with their own written output. Such activities can facilitate their development as proficient and knowledgeable writers and corpus users, which can later assist them in constructing their own corpora and conducting research.

I acknowledge several limitations in this study. Firstly, there was a discrepancy in the size of the corpora, suggesting a need for an increase in the size of the reference corpus. To address this issue, the research findings were normalised to occurrences per 10,000 words. Additionally, there is potential for further research, including the addition of another reference corpus, such as one made up of PhD dissertations written in English (L2) by Czech university students, to investigate developmental patterns. Further research could also involve a broader range of disciplines, possibly including the hard sciences. Despite these limitations, the corpus-based analysis provided valuable insights into how Czech students use endophoric markers in their MTs across three disciplines in the humanities.

Acknowledgements

Works Cited

ABDOLLAHZADEH, Esmaeel. 2019. "A Coss-cultural Study of Hedging in Discussion Sections by Junior and Senior Academic Writers". *Ibérica* 38: 177-202.

ÄDEL, Annelie. 2006. *Metadiscourse in L1 and L2 English*. Amsterdam: John Benjamins Publishing. <https://doi.org/10.1075/scl.24>.

AFZAAL, Muhammad, Muhammad Ilyas CHISHTI, Chao LIU and Chenxia ZHANG. 2021. "Metadiscourse in Chinese and American Graduate Dissertation Introductions". *Cogent Arts & Humanities*: 8 (1). <https://doi.org/10.1080/23311983.2021.1970879>.

BUNTON, David. 1999. "The Use of Higher Level Metatext in Ph.D. Theses". *English for Specific Purposes* 18: 41-56. <https://doi.org/10.1016/S0889-4906(98)00022-2>.

BURNEIKAITĖ, Nida. 2008. "Metadiscourse in Linguistics Master's Theses in English L1 and L2". *Kalbotyra* 59 (3): 38-47.

BURNEIKAITĖ, Nida. 2009. "Endophoric Markers in Linguistics Master's Theses in English L1 and L2". *Žmogus ir žodis* 3: 11-16.

CAO, Feng and Guangwei HU. 2014. "Interactive Metadiscourse in Research Articles: A Comparative Study of Paradigmatic and Disciplinary Influences". *Journal of Pragmatics* 66: 15-31. <http://dx.doi.org/10.1016/j.pragma.2014.02.007>.

CELIEŠIENĖ, Vilija and Erika SABULYTE. 2018. "Metadiscourse Markers in Technical Texts". *International Journal of Linguistics, Literature and Culture* 5 (4): 1-14.

CRISMORE, Avon and Rodney FARNSWORTH. 1990. "Metadiscourse in Popular and Professional Science Discourse". In Walter, Nash (ed.) *The Writing Scholar. Studies in Academic Discourse*. Newbury Park, CA: Sage: 119-136.

CRISMORE, Avon, Raija MARKKANEN and Margaret S. STEFFENSEN. 1993. "Metadiscourse in Persuasive Writing: A Study of Texts Written by American and Finnish University Students". *Written Communication* 10 (1): 39-71. <https://doi.org/10.1177/0741088393010001002>.

CONNOR, Ulla and Ana I. MORENO. 2005. "Tertium Comparationis: A Vital Component in Contrastive Rhetoric Research". In Bruthiaux, Paul, Dwight Atkinson, William Eggington, William Grabe and Vaidehi Ramanathan (eds.) *Directions in Applied Linguistics: Essays in Honor of Robert B. Kaplan*. Bristol, Blue Ridge Summit: Multilingual Matters: 153-164. <https://doi.org/10.21832/9781853598500-015>.

DAHL, Trine. 2004. "Textual Metadiscourse in Research Articles: A Marker of National Culture or of Academic Discipline?" *Journal of Pragmatics* 36: 1807-1825. <https://doi.org/10.1016/j.pragma.2004.05.004>.

DEL SAZ RUBIO, M. Milagros. 2011. "A Pragmatic Approach to the Macro-structure and

Metadiscoursal Features of Research Article Introductions in the Field of Agricultural Sciences". *English for Specific Purposes* 30 (4): 258-271. <https://doi.org/10.1016/j.esp.2011.03.002 >.

Dontcheva-Navrátilová, Olga. 2023. "Self-mention in L2 (Czech) Learner Academic Discourse: Realisations, Functions and Distribution across Master's Theses". *Journal of English for Academic Purposes* 64: 1475-1585. <https://doi.org/10.1016/j.jeap.2023.101272>.

El-Dakhs, Dina Abdel Salam. 2018. "Why Are Abstracts in PhD Theses and Research Articles Different? A Genre-specific Perspective". *Journal of English for Academic Purposes* 36: 48-60. <https://doi.org/10.1016/j.jeap.2018.09.005>.

Flowerdew, Lynne. 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, Tereza. 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>.

Hyland, Ken. 1999. "Talking to Students: Metadiscourse in Introductory Coursebooks". *English for Specific Purposes* 18 (1): 3-26. <https://doi.org/10.1016/S0889-4906(97)00025-2>.

Hyland, Ken. 2004. "Disciplinary Interactions: Metadiscourse in L2 Postgraduate Writing". *Journal of Second Language Writing* 13: 133-151. <https://doi.org/10.1016/j.jslw.2004.02.001 >.

Hyland, Ken. 2005. *Metadiscourse*. London: Continuum.

Hyland, Ken. 2008. "Academic Clusters: Text Patterning in Published and Postgraduate Writing". *International Journal of Applied Linguistics* 18: 41-62.

Hyland, Ken and Feng Kevin Jiang. 2020. "Text-organizing Metadiscourse: Tracking Changes in Rhetorical Persuasion". *Journal of Historical Pragmatics* 21 (1): 137-164. <https://doi.org/10.1075/jhp.00039.hyl>.

Hyland, Ken and Polly Tse. 2004. "Metadiscourse in Academic Writing: A Reappraisal". *Applied Linguistics* 25 (2): 156-177. <https://doi.org/10.1093/applin/25.2.156>.

Jančaříková, Renata. 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>.

Kawase, Tomoyuki. 2015. "Metadiscourse in the Introductions of PhD Theses and Research Articles". *Journal of English for Academic Purposes* 20: 114-124. <http://dx.doi.org/10.1016/j.jeap.2015.08.006>.

Khalili, Assef and Majid Aslanabadi. 2014. "The Use of Metadiscourse Devices by Non-native Speakers in Research Articles". *Journal of Applied Linguistics and Discourse Analysis* 2 (2): 21-34. <https://doi.org/10.22049/jalda.2017.13648>

Khedri, Mohsen, Chan Swee Heng and Seyed Foad Ebrahimi. 2013. "An Exploration of Interactive Metadiscourse Markers in Academic Research Article Abstracts in Two Disciplines". *Discourse Studies* 15 (3): 319-331. <https://doi.org/10.1177/1461445613480588>.

Kilgarriff, Adam, Pavel Rychly, Pavel Smrz and David Tugwell. 2004. "ITRI-04-08 the Sketch Engine". Lexical Computing. Available from <https://www.sketchengine.eu/>.

Kim, Loi Chek and Jason Miin-Hwa Lim. 2013. "Metadiscourse in English and Chinese Research Article Introductions". *Discourse Studies* 15 (2): 129-146. <https://doi.org/10.1177/1461445612471476>.

Kobayashi, Yuichiro. 2017. "Developmental Patterns of Metadiscourse in Second Language Writing". *Journal of Pan-Pacific Association of Applied Linguistics* 21 (2): 41-54. <https://doi.org/10.25256/PAAL.21.2.3>.

Koutsantoni, Dimitra. 2006. "Rhetorical Strategies in Engineering Research Articles and Research Theses: Advanced Academic Literacy and Relations of Power". *Journal of English for Academic Purposes* 5: 19-36. <https://doi.org/10.1016/j.jeap.2005.11.002>.

Lee, Joseph J. and J. Elliott Casal. 2014. "Metadiscourse in Results and Discussion Chapters: A Cross-linguistic Analysis of English and Spanish Thesis Writers in Engineering". *System* 46: 39-54. <http://dx.doi.org/10.1016/j.system.2014.07.009>.

Liao, Jianling. 2020. "Metadiscourse, Cohesion, and Engagement in L2 Written Discourse". *Languages* 25 (5): 1-21. <https://doi.org/10.3390/languages5020025>.

Martín-Laguna, Sofía and Eva Alcón. 2015. "Do Learners Rely on Metadiscourse Markers? An Exploratory Study in English, Catalan and Spanish". *Social and Behavioral Sciences* 173: 85-92. <https://doi.org/10.1016/j.sbspro.2015.02.035>.

Mauranen, Anna. 1993. *Cultural Differences in Academic Rhetoric: A Textlinguistic Study*. Frankfurt am Main: Peter Lang.

Moreno, Ana I. 2008. "The Importance of Comparable Corpora in Cross-cultural Studies". In Connor, Ulla, Ed Nagelhout and William V. Rozycki (eds.) *Contrastive Rhetoric: Reaching to Intercultural Rhetoric*. Amsterdam: John Benjamins: 25-41.

Mu, Congjun, Lawrence Jun Zhang, John Ehrich and Huaqing Hong. 2015. "The Use of Metadiscourse for Knowledge Construction in Chinese and English Research Articles". *Journal of English for Academic Purposes* 20: 135-148. <http://dx.doi.org/10.1016/j.jeap.2015.09.003>.

Mur-Dueñas, Pilar. 2011. "An Intercultural Analysis of Metadiscourse Features in Research Articles Written in English and in Spanish". *Journal of Pragmatics* 43: 3068-3079. <https://doi.org/10.1016/j.pragma.2011.05.002>.

Paltridge, Brian. 2002. "Thesis and Dissertation Writing: An Examination of Published Advice and Actual Practice". *English for Specific Purposes* 16: 61-70. <https://doi.org/10.1016/S0889-4906(00)00025-9>.

Povolná, Renata. 2012. "Causal and Contrastive Discourse Markers in Novice Academic Writing". *Brno Studies in English* 38 (2): 131-148. <https://doi.org/10.5817/BSE2012-2-8>.

Pujol Dahme, Ana and Moisés Selfa Sastre. 2015. "The Transition from University to Publication: Register and Interactional Metadiscourse Features in Immunology Research Written in Catalan and English". *Ibérica* 30: 155-181.

Qiu, Xuyan and Xiaohao Ma. 2019. "Disciplinary Enculturation and Authorial Stance: Comparison of Stance Features among Master's Dissertations, Doctoral Theses, and Research Articles". *Ibérica* 38: 327-348.
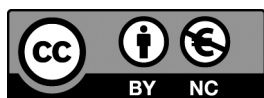
Swales, John. 1990. *Genre Analysis. English in Academic and Research Settings*. Cambridge: Cambridge U.P.

Tribble, Christopher. 2002. "Corpora and Corpus Analysis: New Windows on Academic Writing". In Flowerdew, John (ed.) *Academic Discourse*. London: Longman: 131-149. <https://doi.org/10.4324/9781315838069>.

Vogel, Radek. 2008. "Sentence Linkers in Essays and Papers by Native vs. Non-native Writers". *Discourse and Interaction* 1 (2): 119-126.

40