

Metadiscourse in Lithuanian linguistics research articles: A study of interactive and interactional features

Anna Ruskan

Department of English Philology
Institute of English, Romance and Classical Studies
Faculty of Philology
Vilnius University
Universiteto g. 5
LT-01131 Vilnius, Lithuania
E-mail: anna.ruskan@ff.vu.lt
ORCID iD: 0000-0001-5547-0075

Greta Maslauskienė

Department of English Philology
Institute of English, Romance and Classical Studies
Faculty of Philology
Vilnius University
Universiteto g. 5
LT-01131 Vilnius, Lithuania
E-mail: greta.maslauskiene@ff.vu.lt
ORCID iD: 0000-0002-5319-385X

Abstract. In the recent decades the interpersonal nature of written academic communication has been widely studied in various linguistic/cultural contexts, including Lithuanian. To gain new insights into how knowledge is negotiated interpersonally in Lithuanian scientific texts, the present paper explores the distribution and use of interactive and interactional features of metadiscourse in Lithuanian research articles in a single discipline, i.e., linguistics. For the classification of metadiscourse resources, the study employs the interpersonal model of metadiscourse (Hyland 2005a). Based on the analysis of 30 Lithuanian research articles in the field of linguistics, this exploratory investiga-

Submitted: 05/05/2023. **Accepted:** 09/11/2023

Copyright © 2023 Anna Ruskan, Greta Maslauskienė. Published by Vilnius University Press

This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

tion reveals rhetorical strategies utilized by professional authors to construct a coherent text and engage their audiences in the chosen disciplinary domain. The dominance of interactive metadiscourse features over interactional ones in the corpus examined shows that Lithuanian authors of research articles in linguistics are generally more concerned with organizing discourse and guiding readers through the text than with expressing attitudes and commitment to their arguments. The most common interactive resources are transitions and evidentials, and the most frequent interactional devices include boosters, engagement markers, and hedges. The analysis offers a number of methodological steps necessary for applying the interpersonal model of metadiscourse to Lithuanian data, and complements numerous investigations into Lithuanian academic discourse by illustrating and discussing the writer-reader interaction in linguistics in a larger corpus.

Key words: metadiscourse, interactive, interactional, linguistics, research article, Lithuanian

1 Introduction

Metadiscourse (henceforth MD) is one of the most widely and thoroughly researched areas in the field of academic discourse, as it plays a crucial role in constructing a coherent text, projecting a salient authorial persona, and engaging the reader in the text. MD resources strategically embedded in the propositional content can considerably facilitate the construction of authorial argumentation and establish a dialogic relationship between the writer and the reader that highlights the dynamic and interactive nature of scholarly communication. As claimed by Hyland and Jiang (2018, 19), “with postmodern exceptions, a text communicates effectively only when the writer has correctly assessed both the reader’s resources for interpreting it and his or her likely response to it”. The writer should be aware of any potential objections from the reader, consider their knowledge, anticipate their expectations, and guide them through the processing of the text (*ibid.*, 19). The extensive literature on MD reveals that the writer-reader relation seems to be central to the successful dissemination and negotiation of knowledge in various disciplinary and linguistic/cultural contexts.

The conceptual diversity and complexity of MD features as well as their various means of expression has led to the development of a number of frameworks and typologies (cf. Vande Kopple 1985; Crismore et al. 1993; Ädel 2010). However, one of the most comprehensive approaches to exploring the writer-reader relation in academic texts is the interpersonal model of MD pakeisti į (Hyland 2005a; 2005b), which rests on a basic distinction between ‘interactive’ and ‘interactional’ dimensions. This distinction is based on an earlier classification of interactive (i.e., reader-friendly) and interactional (i.e., reader-managing) devices in discourse proposed by Thompson and Thetela (1995). Interactive devices contribute to the coherent organization of the propositional content

and enhance the reader's comprehension of the text, while interactional devices reflect the author's stance on the information conveyed in the proposition and interaction with the reader (Hyland 2005a, 50–52). The interpersonal model of MD has proved popular in empirical investigations of academic discourse, and is widely characterized as “clear and inclusive [...] building on previous taxonomies” (Abdi et al. 2010, 1671) and as providing “a robust analytic framework” (Lee & Subtirelu 2015, 53).

This approach to MD has been especially effective in revealing how writers guide readers through their arguments and maintain dialogue with them across different academic genres, disciplines, languages and cultures. MD features have been thoroughly examined in research articles, popular science articles, introductory textbooks (Hyland 2005a), book reviews (Tse & Hyland 2006), and websites of research projects (Mur-Dueñas 2021b). Their use in various disciplines, such as business management (Mur-Dueñas 2007; 2011; 2021a), applied linguistics and language teaching (Sheldon 2009), environmental engineering and chemistry (Khedri 2016), medicine (Donadio & Passariello 2022), highlights their role in shaping particular disciplinary cultures. Although the recent decades have seen the extension of MD studies to more languages and cultures (e.g. Estonian, Latvian, Persian, Arabic), considerable attention has been devoted to the exploration of interpersonal relationships in academic texts written in English (Hyland 2005a; 2005b; Ädel 2010; Hyland & Jiang 2016; 2018). As a result, the Anglo-American tradition of writing often becomes a point of departure for analyzing MD in other languages and socio-cultural contexts (Ruskan et al. 2023).

The bulk of research into MD deals with cross-disciplinary and cross-linguistic/cultural variation of MD resources, attributed to the disciplinary conventions of academic texts or language/culture specific factors. For instance, interactional MD devices, such as boosters (Vázquez & Giner 2009), hedges (Šinkūnienė 2012), self-mentions (Khedri 2016), as well as interactive markers, such as transitions and evidentials (Hyland 2010), are more frequent in the ‘soft’ disciplines of the humanities and social sciences than in the ‘hard’ disciplines of science and technology. These distributional differences could be explained by the fact that the soft disciplines are more discursive and driven by interpretation, whereas the hard sciences are usually more based on empirical evidence and facts (Hyland 2005a). Cross-cultural analyses of MD resources reveal the prevalence of these rhetorical devices in texts written by Anglo-American authors. For example, research articles produced by American-based business management scholars contain more MD markers than articles authored by Spanish scholars from the same discipline (Mur-Dueñas 2007; 2010; 2011; 2021a). According to Mur-Dueñas (2011, 3075), the socio-cultural context in which texts are published and read (international vs national target audience) as well as the style of writing (reader or writer responsible) are the

main factors determining these differences. The extensive body of cross-disciplinary and cross-linguistic/cultural research into MD provides many insights into how writers address the needs of their readers and employ diverse rhetorical strategies in various disciplinary and linguistic/cultural contexts.

In the context of Lithuanian academic discourse, MD features have also been studied from a cross-disciplinary and cross-linguistic perspective, operationalizing the interpersonal model of MD. Much attention has been devoted to individual MD categories, such as hedges (Šinkūnienė 2011; 2012), personal pronouns (Šinkūnienė 2018), reformulation markers (Šinkūnienė 2019), as well as to an overall comparison of interactive and interactional MD features across four disciplines (linguistics, literature, sociology, and economics) in English and Lithuanian research articles (Šinkūnienė 2014). These studies provide a quite comprehensive picture of the elements of MD in Lithuanian academic texts and illustrate a number of cross-disciplinary and cross-linguistic similarities and differences in shaping authorial argumentation in English and Lithuanian. They address discipline and language-specific factors determining the distributional properties of MD markers and shed light on the traditions of Lithuanian and Anglo-American writing. However, more research is necessary to present an overall picture of MD features in a specific discipline, reflecting conventions and choices that professional writers of a specific discipline follow in order to construct a coherent text and engage their audience.

This study aims to explore the distribution and use of interactive and interactional MD devices in linguistics, in a larger self-compiled corpus of linguistics research articles by applying the interpersonal model of MD (Hyland 2005a). The methodology for carrying out the study was designed by a research group investigating the features of academic discourse in Estonian, Latvian, and Lithuanian, which represent the dominant national languages of the Baltic states. Thus, the research into MD features in Lithuanian is part of a larger-scale study into academic discourse, which aims to reveal similarities and differences attested in writing traditions in the dominant languages of the Baltic states. The MD features identified in Lithuanian linguistics research articles on the basis of unified methodological procedures will pave the way for comparisons of MD elements in Latvian and Estonian and provide insights into writing conventions in linguistics in the three languages.

2 Data and Methods

To investigate the use of MD devices in Lithuanian linguistics research papers, a specialized corpus was compiled. The corpus consists of 30 authentic Lithuanian research

articles representing original research on various linguistic topics. To form a representative sample, the decision was made to extract all the articles from three different well-known refereed journals, i.e., *Kalbotyra*, *Taikomoji kalbotyra*, and *Lietuvių kalba*. Starting with the last issue of 2020¹ from each journal, the compilers worked down in chronological order and took the first ten single-authored research articles that met the criteria of reporting original linguistic research and were written by native speakers of Lithuanian. In addition, the decision was made not to include multiple articles by the same author in order to maintain data variability and reduce the influence of author idiosyncrasy in the sampled texts. All items that did not meet these criteria were excluded from the sample.

Once the suitable texts were selected, they were manually prepared for analysis by removing abstracts (as they may be regarded as a separate genre of academic discourse) as well as other items that do not typically contain markers of authorial stance and engagement (i.e., tables, figures, lists of references, extended quotations, and linguistic examples provided in the text). The decision was made to retain the acknowledgement sections and endnotes/footnotes, as it was presumed that they would contain features of MD due to their communicative goals. The finished corpus consisted of 30 research articles totalling 135,134 words, as shown in Table 1. The total size of the corpus is similar to those used in other analogous investigations (e.g. Mur-Dueñas 2011) and thus is considered appropriate for the representation of a single discipline.

Journal	Word count	Sentence count
<i>Kalbotyra</i>	46,466	1,976
<i>Lietuvių kalba</i>	34,367	1,821
<i>Taikomoji kalbotyra</i>	54,301	2,403
Total	135,134	6,200

Table 1. Corpus characteristics

After sample texts were collected and their content reduced, the next stage, i.e., data analysis, was conducted by operationalizing Hyland's (2005a) MD model of interactive and interactional resources. The linguistic devices were classified according to the interactive/interactional dimension and marker category, i.e., interactive (transitions, frame markers, endophoric markers, evidentials, code glosses) and interactional (hedgies, boosters, attitude markers, engagement markers, self-mentions), as illustrated in Table 2.

¹ There are two articles from the year 2021.

Interactive dimension	Interactional dimension
Transitions <i>additive</i> <i>contrastive</i> <i>consequential</i> (Hyland & Jiang 2018)	Hedges (e.g. <i>may, possible, seem</i>)
Frame markers <i>sequencers</i> <i>labellers</i> <i>goal announcers</i> <i>topic shifters</i> (Hyland 2005a)	Boosters (e.g. <i>show, clearly, confirm</i>)
Endophoric markers <i>previewing, reviewing</i> (Pisanski Peterlin 2005) <i>visuals</i> (Hyland & Jiang 2018) <i>examples</i>	Attitude markers (e.g. <i>right, fortunately</i>)
Evidentials <i>integral, non-integral</i> (Swales 1990)	Engagement markers <i>reader pronouns</i> (e.g. <i>you</i>) <i>directives</i> (e.g. <i>look, consider</i>) <i>questions</i> <i>appeals to shared knowledge</i> (e.g. <i>of course, as known</i>)
Code glosses <i>reformulation, exemplification</i> (Hyland 2007) <i>elaboration</i>	Self-mentions (e.g. <i>I, we</i>)

Table 2. Hyland's (2005a) MD model and its adaptation

Following Hyland and Jiang (2018, 20), transitions in our data were further divided into those that marked additive, contrastive, or consequential relations between main clauses. The markers *ir* 'and' and *arba* 'or' were excluded because, like the English markers *and* and *or*, they frequently display "routine automaticity as default connectors" (Hyland & Jiang 2018, 23). Frame markers were subdivided into sequencers, labellers, goal announcers, and topic shifters (Hyland 2005a). For the further classification of endophoric markers, we used the traditional subcategorization of these items into those denoting previewing and reviewing (see Pisanski Peterlin 2005) alongside a separate subcategory of markers that directed the reader's attention to visual aids (see Hyland & Jiang 2018). In addition, a separate subcategory was created during a pilot investigation to group all endophoric markers that were used to refer to examples provided in the text. Drawing on the classification of evidentials in Swales (1990), they were grouped into integral (i.e., integrated into the sentence) and non-integral evidentials (i.e., not incorporated in the sentence structure but provided in parentheses). Code glosses were subdivided into reformulation and exemplification markers (Hyland 2007). The additional subcat-

egory labelled ‘elaboration’ was created during the pilot analysis, as it became apparent that there were a number of code glosses that did not fall into either the reformulation or exemplification subcategories, but marked the extension of a provided list of items. Hedges, boosters, and attitude markers were further analyzed in terms of their realization and function (i.e., modal verbs, complement-taking predicates (CTPs), adverbials, adjectives, emphaziers). The items falling into the category of ‘self-mention’ were further subdivided according to their syntactic realization, i.e., whether the marker was realized as a personal pronoun, noun phrase, or the category of person was marked on the verb. Engagement markers were classified as reader pronouns, directives, questions, and markers of shared knowledge (Hyland 2005b). The subcategory of personal asides in Hyland’s (2005b) model of engagement was not applied in the current investigation, as it was seen as too vaguely defined.

It should be noted that instances of MD found in quotations or reports of other scholars’ ideas were not coded, as they may be more reflective of the cited author’s stance. In addition, the decision was made not to include frequency adverbs in the annotation since it was rather difficult to distinguish between their propositional and MD uses.

The corpus was manually annotated by two independent coders following a two-step procedure. First, in the pilot investigation stage, the coders analyzed roughly 10% of the sample to assess and refine the inter-rater agreement. In total, three rounds of double coding were conducted in this stage until an acceptable level of inter-rater agreement was reached. To ensure that the reached agreement did not happen by chance, the inter-rater agreement was calculated using Cohen’s kappa inter-rater measure. The unweighted kappa measure was estimated at 0.970, which is interpreted as “almost perfect agreement” (Landis & Koch 1977, 165). After the three rounds of double coding that allowed both authors to refine the linguistic categories under scrutiny and maintain uniformity in the process of annotating data, the manual analysis of the entire corpus of 30 Lithuanian linguistics research articles was carried out. The data set was then checked for coding errors and inconsistencies, making sure that all markers identified performed metadiscoursal functions. After the manual annotation of the corpus was completed, the frequencies and percentages of identified interactional and interactive features were calculated. The following section provides the findings of the investigation.

3 Results and Discussion

In total, 6,200 sentences from the corpus were analyzed, with 8,433 markers identified and further categorized. The findings reveal how Lithuanian researchers in linguistics build and support their arguments and engage with their readers. As shown in Table 3, researchers used interactive features two times more frequently than interactional ones

(67% and 33%, respectively). The most common subcategories in the corpus were transitions and evidentials.

Interactive			Interactional		
	raw fr.	%		raw fr.	%
Transitions	1,867	33	Boosters	847	30
Evidentials	1,815	32	Engagement markers	793	28
Code glosses	996	18	Hedges	781	28
Endophorics	726	13	Attitude markers	322	11
Frame markers	211	4	Self-mentions	75	3
Total	5,615	100%	Total	2,818	100%

Table 3. Interactive and interactional metadiscourse features in the corpus (raw frequency and percentage)

The predominant use of interactive features in the corpus shows that Lithuanian authors of research articles in linguistics are generally more concerned with organizing discourse and guiding readers through the text than with revealing attitudes and expressing commitment to their arguments. Table 3 provides support for this inference as there is a clear predominance of transitions, evidentials, code glosses, and endophorics in the dataset, compared with boosters, hedges, engagement and attitude markers, as well as self-mentions. Significantly higher frequencies in the categories such as transitions, endophorics, and code glosses reveal that Lithuanian authors in the field of linguistics seek to construct discourse coherently and clearly mark connections between arguments to facilitate reader comprehension. A high number of evidentials in the corpus also indicates Lithuanian authors' frequent attempts to enhance persuasiveness in their texts: evidentials are effective means of providing intertextual weight to authors' claims while demonstrating their expertise in the research territory, giving justifications for their statements, and comparing research findings.

The predominance of interactive markers in academic writing within the field of linguistics has also been attested in studies based on other languages, such as English (e.g. Geng & Wei 2023; Park & Lee 2022; Binmahboob 2022), Persian (e.g. Rahimpour & Faghieh 2009), Arabic (Sultan 2011). For example, having analyzed abstracts of 50 English research articles, Geng and Wei (2023) convincingly demonstrate that linguists used interactive resources approximately two times more frequently than interactional markers. Focusing only on discussion sections of linguistics research articles, Sultan (2011) provides sound evidence for the significantly higher estimate of interactive markers in

English and Arabic academic texts. The prevailing dominance of the interactive dimension has also been reported in the analyses of full-length English research articles in applied linguistics authored by Korean (Park & Lee 2022), Saudi and British authors (Binmahboob 2022). As diachronic studies show, there has been a decline in interactional markers and an apparent increase in the use of interactive markers in academic texts from applied linguistics over the last few decades (see Park & Lee 2022; Hyland & Jiang 2018). This tendency potentially suggests that authors in linguistics are more concerned with textual coherence of their articles and less inclined to involve the reader in their texts (Sultan 2011, 36). Moreover, the apparent tendency to predominantly use interactive resources in linguistics might mean that written scientific communication in this disciplinary domain is becoming more audience-responsible, objective and less persuasive (Park & Lee 2022, 5–8).

The distribution of interactional features in our corpus may suggest language and/or culture-specific features. Unlike in Hyland's (2005a, 145) study, which shows the dominance of hedges and attitude markers in the discipline of applied linguistics, in Lithuanian linguistics articles, the most frequent interactional devices are boosters, followed by engagement markers and hedges, attitude markers, and self-mentions. Thus, Lithuanian authors tend "to restrict, or fend off, possible alternative voices, closing them down using boosters to emphasize the strength of the writer's commitment, and thereby convince the reader through the force of the argument" (Hyland 2005a, 146) rather than to present arguments in cautious ways, using hedges and showing openness to other voices. Less frequent use of hedges may be a national feature, for, as Šinkūnienė's cross-linguistic and cross-disciplinary study (2011) into hedges in research articles on linguistics and medicine in English and Lithuanian confirms, Lithuanian linguists use hedges less frequently than do English writers. Analyses of hedges in other disciplines and languages, such as business administration in Spanish (Mur-Dueñas 2021a) and medicine in Italian (Donadio & Passariello 2022), also report lower frequencies of hedges in comparison to English. The fact that Romance languages show more parallels in the use of hedges with Lithuanian than with English seems to be an interesting linguistic matter with possible cultural roots that could be explored in further studies.

3.1 Interactive metadiscourse features

The distribution of the interactive MD features in the corpus is provided in Table 4 and further illustrated and discussed in the following subsections (transitions, evidentials, code glosses, endophoric markers, and frame markers).

Category	Subcategory	%
Transitions	<i>constrastive</i>	56
	<i>consequence</i>	23
	<i>addition</i>	21
Total		100
Evidentials	<i>non-integral</i>	76
	<i>integral</i>	24
Total		100
Code glosses	<i>exemplification</i>	51
	<i>reformulation</i>	30
	<i>elaboration</i>	19
Total		100
Endophoric markers	<i>examples</i>	32
	<i>visuals</i>	29
	<i>reviewing</i>	25
	<i>previewing</i>	14
Total		100
Frame markers	<i>goal announcers</i>	46
	<i>sequencers</i>	30
	<i>labellers</i>	24
	<i>topic shifters</i>	4
Total		100

Table 4. Distribution of interactive MD features in the corpus

3.1.1 Transitions

Transitions are the most frequent interactive markers in the corpus (33% of all interactive resources). The absolute predominance of these MD markers is attested in the analyses of linguistic research articles in other languages too, such as English (Geng & Wei 2023; Binmahboob 2022; Sultan 2011) and Arabic (Sultan 2011). Realized via conjunctions and adverbials, transitions are important means of organizing discourse and managing reader comprehension. Transitions allow authors to explicitly mark contrast (1), addition (2), and consequence (3) in their texts, revealing patterns of their logical thinking (Hyland 2010, 132):

- (1) *Etninė tapatybė kai kurių mokslininkų yra tapatinama su tautine tapatybe, tačiau kai kuriais aspektais jos vis dėlto skiriasi.* (TK_1)
 ‘Ethnic identity is equated with national identity by some scholars, **but** they differ in some aspects.’

- (2) *Viena idioma perteikia Melavimo freimą [...]. **Taip pat** rastas ir pavyzdys (1), kuriame aktualizuojamas Gimimo freimas [...]. (TK_2)*
 ‘One idiom conveys the frame of Lying [...] Example (1) was **also** found in which the Nativity frame is actualized [...].’
- (3) *Anglų kalboje ought yra laikomas beveik ekvivalentišku should (Coates 1983, 246–247), **taigi** norėdami rekomenduoti, siūlyti ar perspėti anglakalbiai autoriai gali rinktis ir vieną, ir kitą veiksmažodį. (K_1)*
 ‘In English, *ought* is considered almost equivalent to *should* (Coates 1983, 246–247), **thus** English-speaking authors can choose either verb to recommend, suggest or warn.’

Such explicit ways of expressing arguments allow for clear audience orientation. The author’s ability to build claims coherently and persuasively is seen as particularly important in the soft science domains (Hyland 2010, 139). In our corpus, transitions most frequently signal contrastive relations (56%), whereas markers of consequence and addition occur significantly less frequently (23% and 21%, respectively).

The most frequent contrastive transitions in our data are *tačiau* ‘however’ (36%), *o* ‘but’ (26%), and *nors* ‘although’ (11%). The close investigation of *tačiau* ‘however’ and *nors* ‘although’ in the corpus reveals their syntactic versatility, which might be interlinked with their predominance in the data set: both markers frequently occur either in the initial or medial sentence positions. The high frequency of *o* ‘but/and’ in the corpus was expected: it is among the most predominant linguistic items in Lithuanian in general (Utka 2009; Bielinskienė 2010) and is highly multifunctional (Bielinskienė 2009, 61). This marker may signal contrast and addition (Ambrazas et al. 2006; Bielinskienė 2009; Šliogerienė et al. 2015). *O* ‘but/and’ is the second most frequent marker in both categories of addition and contrast: it makes up 26% of contrastive and 19% of additive transitions.

- (4) *Dienotvarkės nustatymo teorija daugiau žvelgia į kiekybinę informacijos pateikimą, o žiniasklaidos rėmų – į kokybinius turinio aspektus. (TK_3)*
 ‘Agenda-setting theory focuses more on the quantitative presentation of information, **whereas** media framing focuses on the qualitative aspects of content.’
- (5) *Mokslinėje literatūroje diskurso markeriai apibrėžiami įvairiai, o pati diskurso markerio sąvoka gali būti suprantama ir siauriau, ir plačiau. (K_2)*
 ‘In the scientific literature, discourse markers are defined in various ways, **and** the concept of a discourse marker itself can be understood both narrowly and broadly.’

An interesting tendency detected in our data is the co-occurrence of *nors* ‘although’ with other contrastive transitions, such as *tačiau* ‘however’, e.g.:

(6) *Nors* tekstynas nėra nei lemuotas, nei anotuotas, **tačiau** didelės apimties autentiška kalbinė medžiaga suteikė galimybę tirti realią, o ne idealizuotą prancūzų ir lietuvių kalbų vartoseną ir objektyviai vertinti kalbinę raišką. (K_3)

*‘**Although** the text is neither lexical nor annotated, [**however**] the large volume of authentic linguistic material provided an opportunity to study real, not idealized usage of the French and Lithuanian languages and to objectively assess linguistic expression.’

(7) *Nors* tiesiogiai nepasakoma, **tačiau** galima suprasti, jog ši „reikšmė“ apima ir veiksmažodžio leksinę reikšmę ir jo aspektinį tipą. (K_4)

*‘**Although** it is not directly stated, [**however**] it can be understood that this “meaning” includes both the lexical meaning of the verb and its aspectual type.’

Semantically speaking, in all such cases detected in the corpus the use of additional transitions such as *tačiau* ‘however’ seems to be redundant, as the contrastive link between the segments is already marked by *nors* ‘although’. Whereas such doubling of contrastive transitions seems incorrect in the English translations, it appears to be acceptable in the Lithuanian originals. These observations suggest that Lithuanian authors place emphasis on strengthening their arguments and guiding the reader toward their preferred interpretation. To illustrate, the author in example (6) addresses the limitations of the corpus used and then ‘recovers’ the credibility of the analysis by emphasizing with *tačiau* ‘however’ the benefits of using the dataset. In (7) we see a similar case, as the author highlights the contrastive relationship between the discourse segments and emphasizes the feasibility of the interpretation he/she suggests by incorporating the additional second transition, guiding the reader towards the preferred interpretation and acceptance of the argument the author builds.

To mark consequential relations, authors mostly employ transitions *taigi* ‘thus’ (44%), *todėl* ‘so’ (35%), and *tad* ‘so’ (11%). Transitions of consequence are crucial elements in the construction of research articles: they enable authors to effectively regulate reader comprehension by directly providing interpretations and drawing implications regarding the obtained results (Hyland & Jiang 2018, 23). This is further attested by looking at the distribution of MD resources across the major structural parts of the research articles in our corpus, as most transitions marking consequence in our data are found in the discussion and conclusions sections (63%).

In addition to the already mentioned *o* ‘and’ (19%), the predominant transitions of addition in our corpus are *taip pat* ‘also’ (45%) and *be to* ‘in addition’ (18%). Mostly used sentence-initially, they represent an important means of organizing discourse and enhancing the reader’s understanding of the text.

3.1.2 Evidentials

The second most common category of interactive resources in the corpus is evidentials (32% of all interactive markers). Indicating sources of information beyond the current text, evidentials are not only used to give credit to other scholars' ideas and establish research territory but also serve as intertextual reinforcement to the author's claims (Hyland 2010, 140). The latter is especially useful in soft disciplines, such as linguistics, which are characterized by a need for a higher degree of persuasiveness (*ibid.*).

Our findings suggest a clear preference for the use of non-integral evidentials (76%) (example 8) as opposed to integral (24%) (9). The most common patterns of integral evidentials were *pasak X* 'according to X', *pagal X* 'according to X', and *remiantis X* 'according to X' (27%, 15%, and 12% of all integral evidentials, respectively).

(8) *Klaipėdos krašto bei vakarų aukštaičių tarmėse daugelis prielinksninių konstrukcijų įsigalėjo dėl vokiečių kalbos įtakos (Laigonaitė 1957, 33).* (K_5)

'In the dialects of the Klaipėda region and western highlands, many prepositional constructions came into force due to the influence of the German language (Laigonaitė 1957, 33).'

(9) *Pasak Hylando (2006), gretinamoji analizė yra labai tinkamas metodas akademinio diskurso studijoms [...].* (K_6)

'According to Hyland (2006), a comparative analysis is a very suitable method for the study of academic discourse [...].'

In general, the use of non-integral evidentials places less emphasis on the referenced authors and more on the cited ideas (Hyland & Jiang 2018, 25). This might suggest that the alternation between the integral and non-integral evidentials in the academic text could be a conscious choice. The cross-linguistic comparison of evidentials in Lithuanian and English research articles in linguistics, economics, and sociology also convincingly demonstrates the tendency for authors to mostly rely on non-integral forms while citing (Šinkūnienė 2014, 121). This lends support to the idea that certain rhetorical elements, such as evidentials, may be universal features in academic discourse across languages, cultures, and disciplines (*ibid.*).

While investigating non-integral evidentials in academic texts, scholars also explore the phenomenon of multiple citations (see Hyland & Jiang 2018; Šinkūnienė 2014). In our data, the number of instances including more than three references in a single case of non-integral citation was unexpectedly high: 20% of all non-integral evidentials. The largest number of multiple citations was 21, as illustrated below:

- (10) *Lietuvių kalbos intarpinių ir sta kamieno veiksmažodžių formaliųjų ir semantinių santykių regularumas su tos pačios šaknies ia kamieno (rečiau kitų struktūrų) veiksmažodžiais kalbininkų yra seniai pastebėtas ir interpretuojamas gana įvairiai (Endzelīns 1951, 764–765; Skardžius 1943, 463–486; Būga 1959, 431–469; Arumaa 1957; Stang 1942; 1966; Kazlauskas 1968, 319–323; LKG 2, 225; Stepanov 1976; Temčīn 1986; Karaliūnas 1987; Pakalniškienė 1993; Palmaitis 1998; Kaukienė 1994, 2002; Valeckienė 1998, 27–28; Villanueva-Svensson 2010; Arkadiev 2008; 2013; Gorbachov 2007; 2014 ir kt.). (LK_1)*

‘The regularity of the formal and semantic relations of Lithuanian inflectional and *sta* stem verbs with verbs of the same *ia* stem (rarely of other structures) has long been noticed by linguists and interpreted quite differently (Endzelīns 1951, 764–765; Skardžius 1943, 463–486; Būga 1959, 431–469; Arumaa 1957; Stang 1942; 1966; Kazlauskas 1968, 319–323; LKG 2, 225; Stepanov 1976; Temčīn 1986; Karaliūnas 1987; Pakalniškienė 1993; Palmaitis 1998; Kaukienė 1994, 2002; Valeckienė 1998, 27–28; Villanueva-Svensson 2010; Arkadiev 2008; 2013; Gorbachov 2007; 2014, etc.).’

The frequent use of multiple citations in the research articles of Lithuanian linguists reveals their conscious efforts to synthesize information from numerous sources, enriching their academic texts with solid intertextual support.

3.1.3 Code glosses

Code glosses, MD markers that allow authors to rephrase or clarify ideational information (Hyland 2010, 129), make up 18% of all interactive resources in our corpus. The majority of code glosses in our data mark exemplification (51%), whereas markers signalling reformulation (30%) and elaboration (19%) were less frequent.

The most common MD devices marking exemplification were *pavyzdžiui* ‘for example’ (52%) and *pvz.* ‘e.g.’ (37%). Upcoming reformulation was typically signalled by *t. y.* ‘i.e.’ (60%) and *kitaip tariant* ‘in other words’ (21%). Elaboration was predominantly marked by *ir kt.* ‘etc.’ (45%) and *ir pan.* ‘etc.’ (27%).

The inclusion of code glosses shows authors’ attempts to help readers grasp information, as these are effective means of clarifying the presented statements, e.g.:

- (11) *Svarbu pastebėti, kad skirtingos funkcijos kalbose išplėtotos nevienodu laipsniu: pavyzdžiui, anglų, olandų kalbose žiūrėk tipo markeriai kur kas dažnesni reiškiant dėmesio atkreipimo funkciją nei klausyk tipo markeriai. (K_7)*

‘It is important to notice that different functions are developed to a different degree in languages: **for example**, in English and Dutch, *look*-type markers are much

more frequent in denoting the function of attracting attention than listen-type markers.’ (K_8)

- (12) *Be to, šis veiksmažodis vienintelis lietuvių kalboje gali išreikšti „visus būtinybės tipus“ (Holvoet 2007, 159), t. y. jis, skirtingai nuo privalėti ar reikėti, turi ir episteminę, ir neepisteminę reikšmę, taip pat kaip ir should.* (K_9)

‘In addition, this verb is the only one in the Lithuanian language that can express “all types of necessity” (Holvoet 2007, 159), i.e., unlike *must* or *need*, it has both epistemic and non-epistemic meaning, just like *should*.’

- (13) *Taigi iš apibrėžties matyti, kad yra receptyvusis žodynas (žinomi, suprantami žodžiai) ir produkcijos (vartojami žodžiai), **kitaip tariant**, dalį savo žodyno žodžių asmuo supranta išgirdęs ar perskaitęs, bet pats jų neprodukuoja, o kitą dalį žodžių aktyviai produkuoja kalbėdamas ar rašydamas.* (TK_4)

‘Thus, it can be seen from the definition that there is a receptive vocabulary (known, understood words) and production (used words), **in other words**, a person understands part of the words in his vocabulary after hearing or reading, but does not produce them himself, and the other part of the words is actively produced by speaking or writing.’

These typical instances of exemplification (11) and reformulation (12 and 13) reveal authors’ attempts to provide more clarification to facilitate the correct understanding of the ideational content for their audiences. In (11) the author extends the information by providing a clear example that supports and contextualizes the argument provided, whereas in (12) the author presents a more detailed restatement of the argument to potentially avoid any confusion and enhance the reader’s correct interpretation of the concepts under discussion. (13) represents another invaluable function of code glosses which is the clarification of terms that could be at first more difficult to grasp. Looking at the examples, it is clear that authors anticipate which parts of the text might require more specification, elaboration, and definition. With the help of code glosses and specifications that follow, they make this information more readily available to their readers.

3.1.4 Endophoric markers

Endophoric markers constitute 13% of interactive markers in the corpus. The most prevalent subcategory consists of markers directing readers’ attention to various examples in the text, amounting to 32% of all endophorics. The most typical markers referring to examples were *X pavyzdyje/pavyzdyje X* ‘in example X’ (37%). The second most frequent subcategory of endophorics (29%) in our data contains markers referring to various visual aids, such as tables, graphs, pictures, etc. The typical markers include *X lentel** ‘Table X’ (30%), *X pav.* ‘in Fig. X’ (27%), *X paveiksl** ‘Figure X’ (15%).

Reviewing endophorics (used to refer to the previously mentioned information in the text), as well as previewing markers (directing readers' attention to upcoming items in the text) were used less frequently (25% and 14%, respectively). The most typical reviewing endophoric markers (45% of all reviewing endophorics) are expressions containing the word *minėti* 'mention' (*kaip minėta* 'as mentioned', *jau minėta* 'already mentioned', *kaip jau buvo minėta* 'as already mentioned', *minėta/minėta/minėtas/minėtu* 'mentioned', etc.). This is directly in line with the findings reported by Šinkūnienė (2014, 86), who investigated endophoric markers in Lithuanian research articles from humanities and social sciences. As our data reveals, these expressions frequently collocate with the adverbial *anksčiau* 'earlier' and *aukščiau* 'above'. Other expressions with more specific reference to particular parts of the text, such as *straipsnio pradžioje* 'at the beginning of the article' (1%), *X skyriuje* 'in section X' (1%) are used less sparingly, possibly due to their less general and versatile character than the aforementioned expressions with *minėti* 'mention'. In a similar vein, previewing endophorics are predominantly (30%) expressed with the marker that conveys very general semantic meaning, i.e., *toliau* 'further'.

Used as important cues that establish connections between various elements in the text, endophoric markers clarify exposition and allow readers to follow the author's argumentation in a coherent manner, as seen below:

- (14) *Akivaizdu, kad pavyzdyje (13a) konceptualioji reikšmė, koduojama tiek veiksmazodyje devoir 'privalėti', tiek dire 'sakyti', yra išblukusi – būtent tai ir yra gramatiškėjimo arba pragmatiškėjimo procesų rodiklis [...]* (K_10)

'It is clear that **in example (13a)** the conceptual meaning encoded in both the verb *devoir* 'to oblige' and *dire* 'to say' is faded; this is exactly the indicator of grammaticalization or pragmaticalization processes [...].'

- (15) *Vis dėlto, kaip jau minėta įvade, šios konstrukcijos negali būti laikomos tikrais rezultatyvais, nes antriniu predikatu išreikštas požymis realizuojasi nepasibaigus veiksmui, o jo metu.* (K_11)

'However, **as already mentioned in the introduction**, these constructions cannot be considered real resultatives, because the attribute expressed by the secondary predicate is not realized after the action, but during it.'

The endophoric marker *pavyzdyje (13a)* 'in example (13a)' in (14) is used to link the proposition the author sets forth with the specific illustration that supports it. Not only does it allow the author to explicate argumentation, but to also effectively guide the audience by reinforcing a correct interpretation of the concepts under discussion. Similarly, with the help of the reviewing marker *kaip jau minėta įvade* 'as already mentioned in the introduction' in (15), the author links the information at hand with what is mentioned in the other part of the text (i.e., in the introduction). Such use of MD elements helps

readers to recall certain arguments provided previously, which in turn helps authors to direct their audiences towards the preferred interpretation and at the same time enhance the overall coherence of the text.

3.1.5 Frame markers

Frame markers are the least frequent MD resources in our corpus, constituting only 4% of all detected interactive devices. They organize discourse by signalling discursive goals and upcoming directions of argumentation (16), sequencing information (17), indicating topic shifts (18), and labelling text stages (19):

- (16) *Šio tyrimo tikslas – išsiaiškinti, ar XIX a. pabaigoje–XX a. pradžioje leistos knygos atspindi jų autorių pasirinktą ortografijos modelį.* (TK_5)
 ‘The **purpose of this research** is to find out whether the books published in the late 19th-early 20th century reflect the orthography model chosen by their authors.’
- (17) *Tyrimas turi tam tikrų ribotumų: pirmas – trumpas egzaminų temų sąrašas, antras – tekstyno dydis, kuris koreliuoja su mokinių, pasirinkusių laikyti prancūzų kalbos egzaminą, skaičiumi.* (K_12)
 ‘The study has some limitations: **first**, the short list of exam topics, and **second**, the size of the corpus, which correlates with the number of students who chose to take the French exam.’
- (18) *Toliau analizuojamas kitų kalbų elementų kiekybinis pasiskirstymas abiejų amžiaus grupių merginų ir vaikinų šnekoje (2 lentelė).* (LK_2)
 ‘**Next**, the quantitative distribution of other language elements is analyzed in the speech of girls and boys of both age groups (Table 2).’
- (19) *Apibendrinant kiekybinio ir kokybinio tyrimo rezultatus, galima daryti toliau pateikiamas pagrindines išvadas.* (LK_3)
 ‘**Summarizing** the results of the quantitative and qualitative research, **the following main conclusions can be drawn.**’

As seen from the examples above, frame markers are powerful rhetorical resources that bring clarity to exposition and allow authors to make scientific information more manageable and accessible to their readers.

The most frequent type of frame markers are goal announcers (46%), predominantly expressed by *tyrimo/straipsnio tikslas* ‘the purpose of the study/research paper’ and *siekiamo* ‘aim to’, which comprised 28% and 20% of all goal announcers, respectively. Sequencing markers constitute the second most common subcategory of frame markers (30%). Typically, they are realized via adjectival/adverbial phrases (*visų*) *pirma* ‘first of all’ (23%), *añtra* ‘second’ (15%), *pirmiausia* ‘first of all’ (11%), *galiausiai* ‘finally’ (11%). Markers

used to label text stages and rhetorical functions make up 24% of all frame markers. Labelers predominantly mark the drawing of conclusions, a tendency also detected in English research articles (Hyland & Zou 2020, 38). They are mostly expressed by *apibendrinant/apibendrinus* ‘in summary’ (42%) and *apskritai* ‘in general’ (20%). Topic shift makes up the least frequent category (4%) and is primarily realized via *kitas* ‘next’ (66%).

3.2 Interactional metadiscourse features

The distribution of the interactional MD features in the corpus is provided in Table 5 and further illustrated and discussed in the following subsections (boosters, engagement markers, hedges, attitude markers, and self-mentions).

Category	Subcategory	%
Boosters	<i>lexical verbs of showing, seeing, confirming</i>	60
	<i>emphasizers</i>	19
	<i>adjectives and adverbials</i>	11
	<i>particles and their combinations, modal verbs, nouns and other markers</i>	10
Total		100
Engagement markers	<i>directives</i>	78
	<i>shared knowledge</i>	15
	<i>reader mentions</i>	4
	<i>questions and other markers</i>	3
Total		100
Hedges	<i>modals</i>	39
	<i>adverbials</i>	22
	<i>CTPs</i>	18
	<i>diminishers</i>	10
	<i>approximators</i>	6
	<i>other markers</i>	5
Total		100
Attitude markers	<i>adjectives</i>	87
	<i>adverbs</i>	9
	<i>verbs and other markers</i>	4
Total		100
Self-mentions	<i>first-person singular verb inflection</i>	42
	<i>the author</i>	33
	<i>first-person plural verb inflection</i>	13
	<i>pronouns</i>	12
Total		100

Table 5. Distribution of interactional MD features in the corpus

3.2.1 Boosters

Boosters are the most common type of interactional MD category, accounting for 30% of the interactional features in our corpus. The dominance of boosters points to the Lithuanian linguists' convincing argumentation, persuasion, and high degree of commitment to reported information. As shown in a cross-disciplinary study by Skorczynska and Carrió-Pastor (2021, 584) into verb boosters, these interactional devices are indeed common in research articles in linguistics, compared to their lower frequencies in engineering or medicine. Similarly, a cross-disciplinary and cross-linguistic study into MD features in English and Lithuanian (Šinkūnienė 2014, 96) provides evidence that boosters are more frequent in the field of linguistics in Lithuanian compared to economics or sociology.

More than half of the boosters (60%) in the articles investigated are expressed by lexical verbs of showing, indicating, revealing (e.g. *(pa)rodyti* 'show', *atskleisti* 'reveal'), seeing (e.g. *matyti* 'see', *(pa)stebėti* 'notice'), and confirming (e.g. *patvirtinti* 'confirm', *liudyti* 'witness'), which allow authors to present the results of their studies and report on studies carried out by other scholars in a confident manner (cf. Hyland 2005a, 79), e.g.:

(20) *Dialektometrijos metodais atlikta analizė įrodo visų trijų kartų atstovų vartojamo tarminio varianto homogeniškumą.* (LK_4)

'The analysis conducted by means of dialectometrics **proves** that representatives of all three generations show a homogeneous use of this dialectal variant.'

(21) *Matyti, kad minėtų tyrimų rezultatai kiek kontraversiški, nepateikiantys vieningo atsakymo apie lietuvių kalbos atriciją diasporoje.* (TK_6)

'It can be **seen** that the results of the mentioned studies are somewhat contradictory and do not provide a unified answer regarding the attrition of the Lithuanian language in the diaspora.'

The verb *įrodo* 'proves' (20) collocates with the inanimate subject *analizė* 'analysis' and presents strong authorial arguments, eliminating alternative interpretations. The impersonal form *matyti* 'see/can be seen' (21) contributes to the shared status of the argument, for the available evidence that this verb refers to is accessible to both the author and the reader.

The second most frequent type of boosters found in the articles analyzed is realized by a number of emphasizees (19%). The most common emphasizees functioning as boosters are *ypač* 'especially' (38%), *būtent* 'exactly, precisely' (25%), *itin* 'particularly' (19%), *gerokai* 'significantly, considerably' (10%), which amplify the author's argumentation and express a high degree of conviction, e.g.:

- (22) *Būtent* Aleksa Girdenio ir Zigma Zinkevičiaus lietuvių tarmių klasifikacijoje pagal šį požymį skiriamos šiaurės žemaičių patarmės. (LK_5)
 ‘It is in Aleksas Girdenis’ and Zigmą Zinkevičius’ classification of Lithuanian dialects that these subdialects of northern Lowland Lithuanian are identified according to this feature.’
- (23) *Ypač* jei namie mažumos kalba ne viena [...] gali būti nutarta, kad pakanka, jei vaikas jomis šneka. (TK_7)
 ‘Especially if there is more than one minority language spoken at home [...], a decision could be made that it is sufficient if a child speaks these languages.’

The emphasizer *būtent* ‘exactly, precisely’ (22) highlights how, in the classifications of Lithuanian dialects designed by two renowned Lithuanian scholars, Lowland Lithuanian subdialects are identified on the basis of a particular feature. Similarly, *ypač* ‘especially’ (23) strengthens the author’s argument that if a family speaks more than one minority language, parents can decide that their children develop only speaking skills of this language.

11% of the boosters in the corpus are expressed by epistemic adjectives and adverbials. The most frequent epistemic markers conveying a high degree of commitment are *aišku* ‘clear’, *akivaizdu* ‘evident’, *aiškiai* ‘clearly’, *akivaizdžiai* ‘evidently’, *iš tiesų* ‘in fact, really’, *tikrai* ‘really’, *be abejo* ‘no doubt’, and *neabejotinai* ‘undoubtedly’. Boosters are least frequently realized by particles and their combinations, modal verbs, nouns, and other means (10%).

3.2.2 Engagement markers

Markers of engagement, alongside hedges, are the second most frequent category, comprising 28% of all interactional devices in the corpus. The most frequent engagement resources are directives (78%), followed by appeals to shared knowledge (15%), reader mentions (4%), questions (2%), and other markers (1%). Directives are a predominant subcategory, as their most frequent expressions are imperative forms such as *žr.* ‘see’ (44%) and *plg.* ‘cf.’ (29%), which guide the reader through the sections of the text, draw attention to visual data, or indicate intertextual links. Alongside imperatives, we have identified participles of necessity (10%), modal verbs of necessity (9%), as well as adjectives of necessity and importance (7%), which also represent directives (cf. Hyland & Jiang 2016), as illustrated below:

- (24) *Reikia* paminėti ir kai kuriuos apklausos ribotumus. (TK_8)
 ‘Some limitations of the survey **should** be mentioned.’
- (25) *Pastebėtina*, kad apie sėkmingą savo vaikų dvikalbę raidą tėvai noriai kalbasi su aplinkiniais, [...]. (TK_9)
 ‘It **should be noted** that parents willingly share their experience with others on how their children successfully developed bilingualism [...].’

(26) *Svarbu prabrėžti, kad dalis respondentų (37 asmenys) mokėsi ar studijavo ne tik Lietuvoje, bet ir Airijoje (žr. 2 lentelę). (TK_10)*

‘It is **important** to emphasize that part of the respondents (37 people) studied not only in Lithuania but also in Ireland (see Table 2).’

The modal verb *reikia* ‘be necessary’ (24) collocating with the speech act verb *paminėti* ‘to mention’ draws the reader’s attention to some limitations of the study. In a similar vein, the impersonal participle *pastebėtina* ‘should be noted’ (25) and the adjective *svarbu* ‘important’² collocating with the speech act verb *prabrėžti* ‘to emphasize’ (26) encourage the reader to consider the author’s observations and claims. The participles of necessity also derive from speech act verbs (i.e., *pažymėti* ‘to note’ > *pažymėtina* ‘should be noted’, *konstatuoti* ‘state’ > *konstatuotina* ‘should be stated’), mental verbs (i.e., *pastebėtina* ‘to observe’ > *pastebėti* ‘should be observed’, *prisiminti* ‘to remember’ > *prisimintina* ‘should be remembered’), or the verb that explicitly refers to the reader’s attention (*atkreipti dėmesį* ‘to draw attention’ > *atkreiptinas dėmesys* ‘attention should be drawn’). These markers allow authors to emphasize the importance of some phenomena and involve the reader in noticing, marking, and considering arguments.

Appeals to shared knowledge, by means of which “writers construct themselves and their reader as members of the same academic community” (Hyland & Jiang 2016, 35), are expressed in our corpus most frequently by participle or adjective-based CTPs or parentheticals *žinoma* ‘(is) known to, of course’, *manoma* ‘(is) thought to/reportedly’, *suprantama* ‘(is) understood to/understandably’, *natūralu* ‘natural/naturally’, which make up 65% of this feature. Appeals to shared knowledge are also realized by markers that refer to the reader’s knowledge of research processes/practices, for example, *paprastai* ‘typically’, *tipinis* ‘typical’, or those that allude to the reader’s “assumed familiarity with aspects of background information” (Hyland & Jiang 2016, 36), such as *tradiciškai* ‘traditionally’, *įprasta* ‘usual’, which make up 35% of appeals to shared knowledge. Indications of shared knowledge help the reader navigate through both old and new information, as illustrated below:

(27) *Kalbinė logika ir intuicija tokiems samprotavimams, žinoma, prieštarauja. (LK_6)*
‘Language logic and intuition, **of course**, contradict such reasoning.’

² It should be noted that adjectives of importance, as in example (26), could be ambiguous between engagement and attitude markers. Occurring in the construction with speech act verbs, they encourage the reader to pay attention to important facts and thus function as engagement devices. However, the inherent semantic meaning of importance that these adjectives convey indicates their functional connection with attitude markers, which qualify various entities as important, significant, relevant, or interesting. Despite the possible ambiguity, we decided to classify such cases as markers of engagement because in combination with a speech act verb, the adjective draws the reader’s attention to the following propositional content.

- (28) *Čia galima išvesti kad ir tokią paralelę: gerai žinomas dalykas, kad baltų liepa yra antrinis pavadinimas medžiui įvardinti, pirminis buvo lenta.* (K_13)
‘We can draw the following parallel here: it is a **well-known thing** that the Baltic ‘linden tree’ is a secondary name for a tree, the primary one was ‘board’.’

The parenthetical *žinoma* ‘of course’ (27) activates the reader’s knowledge and makes the author’s argumentation more persuasive, for the shared knowledge cannot be disputed as it is obvious or self-evident. References to the reader’s familiarity with background knowledge (28) also serve as effective engagement strategies.

The least frequent features of engagement are reader mentions, expressed by the first person plural forms of the verb (i.e., *apžvelgsime* ‘we will review’, *matome* ‘we can see’, *turime pripažinti* ‘we have to admit’), and questions. By incorporating the reader into the text through the use of the first person plural form of the verb (29), the author helps the reader to understand the arguments. Questions functioning as strategies of engagement stimulate the reader’s interest in the topic and emphasize issues that cause problems, controversies, and which are expected to be addressed by the reader, e.g.:

- (29) *Matome, kad anglų kalba matoma ir kaip individų ar jų grupių, ir kaip valstybės tapatybės raiškos išteklius [...].* (TK_11)
‘**We can see** that English is viewed as a resource for expressing individual, group, and state identity [...].’
- (30) *Klausimas tik, kodėl juos tada reikia priskirti rytų baltams?* (K_14)
‘The question is why we have to classify them as Eastern Balts.’

In (30), the question deals with why some features should be attributed to the Eastern Balts if there are arguments against this. Thus, questions may stimulate the reader’s reflection on the argumentation provided by the author.

3.2.3 Hedges

Hedges are as frequent as engagement markers, comprising 28% of all interactional MD features. These findings are not surprising, since hedges, like other interactional devices, are expected to be encountered in discursive disciplines, such as linguistics, where it is important for authors to mark a reduced level of commitment to presented arguments and claims, and to make room for alternative interpretations. As maintained by Šinkūnienė (2012, 138), “with the help of hedges scientists can phrase their claims, report results, draw conclusions and express criticism with caution”.

In the articles examined, hedges are realized by a variety of devices, such as modals (39%), adverbials (22%), CTPs (18%), markers of vague language (16%) (i.e., approx-

imators, diminishers), and other means (5%) (i.e., the subjunctive mood, participles, etc.). The dominance of the modals as hedges can be explained by the high frequency of the impersonal modal *galima*³ ‘it is possible’, which makes up 66% of this subcategory. *Galima* ‘it is possible’ functions as a hedge when it co-occurs with lexical verbs of saying (i.e., *teigti* ‘claim’, *įvardyti* ‘name’, *apibendrinti* ‘generalize’) and cognition (i.e., *manyti* ‘think’, *laikyti* ‘find, hold’, *suprasti* ‘understand’) (cf. Šinkūnienė 2011; 2020), e.g.:

(31) *Galima teigti, kad prieš eidami į mokyklą, vaikai jau geba pasirinkti ir vartoti įvairesnes nurodymų formas.* (LK_7)

‘It is possible (lit. can be claimed) that before going to school, children are already able to choose and use a variety of commands.’

(32) *Kita vertus, galima manyti, jog egzistuoja ir auditorijos suvokimo asimetrija – teigiamais žiniasklaidos pranešimais patikima lėčiau ir sunkiau negu neigiamais, ypač jeigu išgyvenama asmeninė situacija yra sunki.* (TK_12)

‘On the other hand, it is possible (lit. can be thought) that that there is an asymmetry in how the audience perceives media messages. Positive media messages are not trusted as easily as negative ones, especially if a person is experiencing a difficult situation.’

In (31), a tentative conclusion is drawn that children acquire certain linguistic competence (use of commands) before going to school, and in (32), a suggestion is made that the audience does not believe in positive messages conveyed by the media so easily. The examples above illustrate the CTP construction, in which the unit *galima* ‘it is possible’ + infinitive functions as a CTP controlling a *that*-clause, which is the most frequent syntactic pattern of the form *galima* ‘it is possible’ (Šinkūnienė 2020). It should also be noted that the function of *galima* ‘it is possible’ depends on the semantic type of the lexical verb with which it co-occurs (cf. Šinkūnienė 2020, 89). Co-occurring with the lexical verbs *pastebėti* ‘note’, *pridurti* ‘add’, and *pažymėti* ‘mark’, *galima* ‘it is possible’ is used as an engagement marker but not a hedge (ibid.), for the author does not express a cautious opinion but involves the reader to share the author’s observations. The subcategory of modals also includes the 3rd-person *galėti* ‘can, may’ forms, such as *gali* ‘can, may’ (present), *galėjo* ‘could, might’ (past), *galėtų* ‘could, might’ (subjunctive), which make up 34% of all the modals.

The majority of adverbials functioning as hedges express limited authorial knowledge and a reduced degree of commitment to a proposition, i.e., different qualifications of epistemicity (Boye 2012). Among the most frequent adverbials are *gall/galbūt/turbūt/galimai* ‘per-

³ The marker *galima* ‘it is possible’ is a non-agreeing present passive participle.

haps, maybe, possibly' (21%), *greičiausiai/veikiausiai/veikia/ko gero* 'most likely/probably' (18%), which convey the meaning of epistemic possibility or probability. Evidential-epistemic adverbials such as *matyt/regis/atrodo* 'evidently/seemingly/presumably', which make up 17% of all adverbial hedges, acquire a hedging function due to the insufficient perceptual or conceptual evidence that the author has at his/her disposal:

- (33) *Regis, nėra nė vieno akustinio parametro, kuriam vienu ar kitu laikotarpiu nebūtų priskirta galia diferencijuoti kalbamuosius prozodinius elementus.* (LK_8)
 'Seemingly, there is no single acoustic parameter on the basis of which they would not attribute the capacity to differentiate these prosodic elements at some point.'
- (34) *Tarptautinėje arenoje moksliniai ginčai šia tema, atrodo, netyla ir šiandien.* (K_15)
 'In the international arena, these scientific debates, **it seems**, are still relevant today.'

In the examples above, *regis* 'seemingly' and *atrodo* 'seem' make the author's statements less assertive and more cautious, since there is no sufficient justification to make a strong claim or argument. The range of adverbials performing hedging functions in the corpus examined is quite broad. Hedges are also expressed by adverbials such as *iš esmės* 'essentially' (13%), *kone/bene/vargu* 'hardly' (10%), *tarsi* 'as if' (8%), *manytina/spėtina/tikėtina* 'should be thought, should be presumed/likely' (7%), and miscellaneous other markers (1%).

Like adverbial hedges, hedges expressed by verb-based (35), adjective-based (36), and noun-based (37) CTPs controlling *that*-clauses (Usonienė 2013) convey the speaker's assumptions and tentative conclusions about the results of the research:

- (35) *Atrodo, kad mišrios šeimos liuanistinį ugdymą renkasi rečiau [...].* (TK_13)
 'It seems that mixed families choose Lithuanian upbringing for their children less often [...].'
- (36) *Neaišku, kodėl T testas nepatvirtina koreliacijos, bet tikėtina, kad tai būtų galima patikrinti atlikus išsamesnį tyrimą ir surinkus daugiau duomenų.* (TK_14)
 'It is not clear why the T-test does not confirm the correlation, but it is **likely** that checking the validity of the test would be possible after conducting a more comprehensive study and collecting more data.'
- (37) *Tačiau neatmestina prielaida, jog šios konstrukcijos galėjo būti vartojamos gimtosiose vertusiujų šnektose.* (K_16)
 'However, we cannot reject the **assumption** that these constructions could have been used in the native regions of the translators' subdialects.'

(36) and (37) illustrate the co-occurrence of several hedges; alongside the CTPs, in (36) the author's hypothetical interpretation is expressed by the subjunctive form *būtu* 'would be' and in (37) by the modal verb *galėjo* 'could, might'. Common CTPs include such markers as *leidžia manyti/teigti/daryti išvadą* 'allows to believe/claim/draw a conclusion', *manytina* 'should be thought', *spėtina* 'should be guessed', *hipotezė* 'hypothesis', *galimybė* 'possibility', etc.

Among the frequent resources of hedging are also approximators and diminishers. The most common approximator functioning as a hedge is *beveik* 'almost' (68%), followed by *bent kiek* 'at least some' (13%) and *apie* 'about' (11%). The most frequent diminishers are *gana* 'quite' (73%) and *kiek* 'somewhat' (13%). Less frequent are *šiek tiek* 'a little bit' (6%), *iš dalies* 'partly, to some extent' (5%), and *daugiau ar mažiau* 'more or less' (3%). Thus, Lithuanian linguistics articles display a variety of hedging devices that create dialogicity with alternative views and voices, suggesting a cautious presentation of arguments and claims.

3.2.4 Attitude markers

Attitude markers comprise 11% of all interactional resources and are most frequently expressed by evaluative adjectives or adverbs. The low frequencies of attitude markers in our corpus contrast strongly with the findings of other studies. For example, in applied linguistics research articles in English (Hyland 2005a, 145), attitude markers are the second most frequent category. It seems that Lithuanian writers are not as involved in expressing value judgements and attitudes as writers in English, and prefer a more matter-of-fact style of writing, containing less explicit subjectivity.

The most frequent evaluative expressions in the corpus examined are adjectives (87%), followed by adverbs (9%) and verbs/other markers (4%). Common attitudes conveyed by these markers are importance/meaningfulness (28%) and interest (13%), which is not surprising, as in their argumentation scholars have to highlight the role of other studies (38), the importance of processes and phenomena explored (39), and to emphasize the author's interest in the subject matter (40):

(38) *Tam tikrais konkrečiais istorinės sintaksės ir kartu kai kurių linksnių istorinės raidos klausimais yra reikšmingi Vytauto Ambrazo tyrimai.* (K_17)

'Some concrete questions of historical syntax as well as the historical development of particular cases are dealt with in **important** studies by Vytautas Ambrazas.'

(39) *Svarbiausia, kad identifikuojant mažąsias baltų kalbas pasirinkus istorinių regionų kriterijų to kriterijaus nuosekliai nebuvo laikomasi.* (K_18)

'**The most important thing** is that, when identifying small Baltic languages on the basis of the criterion of historical regions, this criterion was not applied consistently.'

- (40) *Idomu, kad iš Paryžiaus jis rašė tik lenkų kalba, o vėliau, jau grįžęs į Lenkiją, iš Varšuvos rašė prancūziškai [...].* (TK_15)
 ‘It is **interesting** that when he was in Paris, he wrote only in Polish; later, when he returned to Poland, from Warsaw he wrote only in French [...].’

It should be noted that values of importance and interest are most commonly identified in the introduction, results and discussion sections, where authors are expected to highlight the assessment of other studies as well as the results of their own research. Various entities in Lithuanian linguistics articles are also discussed in terms of their positivity/negativity (e.g. *tinkamiausias* ‘most suitable’, *neįtikinamai* ‘not convincingly’), comprehensibility (e.g. *aiškiai* ‘clearly’, *išsamiai* ‘comprehensively’), complexity (e.g. *sudėtingas* ‘complex’) and difficulty (e.g. *sunku* ‘difficult’), (un)expectedness (e.g. *nenuostabu* ‘not surprisingly’), and authenticity (e.g. *aktualus* ‘relevant’), which allow writers to express multiple judgements and opinions. The wide range of attitudinal meanings expressed by the author reflects a concern on the part of the author to influence or position the reader’s stance in relation to other studies as well as those conducted by the author himself or herself.

3.2.5 Self-mentions

Self-mentions constitute 3% of all interactional devices in the corpus. The most frequent expressions are verbs used in the first-person singular (42%), followed by references to the author (33%), verbs in the first-person plural (13%), and the pronoun *mano* ‘my’ (12%). These findings are in line with Šinkūnienė’s study (2014, 101), which shows that authorial voice expressed by the pronouns *aš* ‘I’, *mes* ‘we’, or first-person singular or plural verbal forms is not common in Lithuanian articles. These findings are perhaps not surprising, as similar results have been obtained in other studies (i.e., Spanish, Italian, German, French) (ibid., 102). The results in our corpus show that almost all cases of self-mentions are used in one article, which reflects what might be called an individualistic style of writing. The author of the article creates argumentation through the use of self-mentions, e.g.:

- (41) *Šie žingsniai, esu įsitikinęs, puikiai pasitarnautų ne tik mažosioms, bet ir didžiosioms baltų kalboms ir jų istorijai.* (K_19)
 ‘These steps, **I am convinced**, would perfectly serve not only small but also big Baltic languages and their histories.’
- (42) *Atvirai tariant, į perspektyvą įrodyti, kad iš tikrųjų egzistavo mažosios baltų kalbos, asmeniškai žiūriu gana skeptiškai [...].* (K_20)
 ‘Frankly speaking, **I am** rather skeptical about proving that small Baltic languages really existed.’

The use of self-mentions contributes to the persuasiveness of an argument and a salient authorial voice. Their lower frequencies in Lithuanian articles could be connected to different cultural conventions of academic writing. In Lithuanian tradition, the author's voice is expressed through less explicit resources of subjectivity than first-person pronouns or verbs.

4 Conclusions

This exploratory investigation has provided valuable insights into how Lithuanian linguists use MD features to construct their research articles, build their authorial presence, and engage with their audiences. The results of our corpus analysis indicate that Lithuanian authors of research articles in linguistics place a greater emphasis on organizing discourse and guiding readers through the text rather than on expressing their attitudes and commitment to their arguments. This is supported by the higher frequency of MD markers such as transitions, evidentials, code glosses, and endophorics, compared to other categories such as boosters, hedges, engagement markers, attitude markers, and self-mentions. Additionally, the high frequency of evidentials suggests that Lithuanian linguists often attempt to make their written claims more persuasive by demonstrating their expertise in the research field, providing justifications, and comparing findings. Overall, the results suggest that Lithuanian authors in linguistics aim to construct discourse that is coherent and easy for readers to understand, while also seeking to enhance the persuasiveness of their texts.

The present study not only provides an illustration and discussion of MD features in Lithuanian linguistics articles but also offers a number of methodological steps necessary for applying the interpersonal model of MD to other disciplines and languages. Double coding and adapting some MD categories result in an efficient application of the model. Since this research is part of a larger scale study into academic discourse in Estonian, Latvian, and Lithuanian, its results can be compared with those in the other two languages of the Baltic states. Hopefully, the unified methodological approach to exploring MD in the three languages will result in a comprehensive picture of how authors organize linguistic discourse, voice their attitudes, and interact with readers in the three Baltic academic cultures.

Further investigation into MD features in Lithuanian academic texts could be extended by incorporating other genres and discourse types. In addition, future studies could compare MD devices in spoken and written Lithuanian, as there might be apparent differences in the use of certain items, such as transitions regarding their syntactic behaviour and functions in spoken and written texts.

Acknowledgements

This study is supported by the project *Bwrite: Academic Writing in the Baltic States: Rhetorical Structures through culture(s) and languages* EMP475, funded by Iceland, Liechtenstein and Norway through the EEA Grants and Norway Grants. We are sincerely grateful to Nicholas Groom for reading our manuscript and offering his helpful comments.

References

- Abdi, Reza, Manoochehr Tavangar Rizi & Mansoor Tavakoli. 2010. The cooperative principle in discourse communities and genres: A framework for the use of metadiscourse. *Journal of Pragmatics* 42 (6), 1669–1679. <https://doi.org/10.1016/j.pragma.2009.11.001>
- Ädel, Annelie. 2010. Just to give you kind of a map of where we are going: A taxonomy of metadiscourse in spoken and written academic English. *Nordic Journal of English Studies* 9 (2), 69–97. <http://dx.doi.org/10.35360/njes.218>
- Ambrasas, Vytautas, ed. 2006. *Dabartinės lietuvių kalbos gramatika*. [Grammar of Modern Lithuanian]. Vilnius: Mokslo ir enciklopedijų leidybos institutas.
- Bielinskienė, Agnė. 2009. Relevancijos teorija ir diskurso jungtukų bei jungiamųjų žodžių pragmatika. [Relevance theory and pragmatics of discourse connectors and connective words]. *Kalbų studijos* 15, 53–62.
- Bielinskienė, Agnė. 2010. *Sujungiamojo ryšio semantika, pragmatika ir vartosenos ypatumai*. [Semantics, pragmatics and peculiarities of usage of coordination]. Doctoral dissertation. Kaunas: VDU.
- Binmahboob, Thamer. 2022. The Use of Metadiscourse by Saudi and British Authors: A Focus on Applied Linguistics Discipline. *English Language Teaching* 15 (2), 78–89. <https://doi.org/10.5539/elt.v15n2p78>
- Boye, Kasper. 2012. *Epistemic meaning: a crosslinguistic and functional-cognitive study*. Berlin: de Gruyter Mouton.
- Crismore, Avon, Raija Markkanen & 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.
- Donadio, Paolo & Mattia Passariello. 2022. Hedges and boosters in English and Italian medical research articles: A cross-cultural comparison. *International Journal of Language Studies* 16 (1), 1–20.
- Geng, Hui & Han Wei. 2023. Metadiscourse Markers in Abstracts of Linguistics and Literature Research Articles from Scopus-Indexed Journals. *Journal of Modern Languages* 33 (1), 29–49. <http://dx.doi.org/10.22452/jml.vol33no1.2>

- Hyland, Ken. 2005a. *Metadiscourse: Exploring Interaction in Writing*. London/New York: Continuum.
- Hyland, Ken. 2005b. Stance and engagement: a model of interaction in academic discourse. *Discourse Studies* 7 (2), 173–192. <https://doi.org/10.1177/1461445605050365>
- Hyland, Ken. 2007. Applying a gloss: Exemplifying and reformulating in academic discourse. *Applied linguistics* 28 (2), 266–285.
- Hyland, Ken. 2010. Metadiscourse: Mapping interactions in academic writing. *Nordic Journal of English Studies* 9 (2), 125–143. <https://doi.org/10.35360/njes.220>
- Hyland, Ken & Feng (Kevin) Jiang. 2016. “We must conclude that”: A diachronic study of academic engagement. *Journal of English for Academic Purposes* 24, 29–42.
- Hyland, Ken & Feng (Kevin) Jiang. 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, Ken & Hang (Joanna) Zou. 2020. In the frame: Signalling structure in academic articles and blogs. *Journal of Pragmatics* 165, 31–44. <https://doi.org/10.1016/j.pragma.2020.05.002>
- Khedri, Mohsen. 2016. Are we visible? An interdisciplinary data-based study of self-mention in research articles. *Poznań Studies in Contemporary Linguistics* 52 (3), 403–430. <https://doi.org/10.1515/psicl-2016-0017>
- Landis, J. Richard & Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33 (1), 159–174.
- Lee, Joseph J. & Nicholas C. Subtirelu. 2015. Metadiscourse in the classroom: A comparative analysis of EAP lessons and university lectures. *English for Specific Purposes* 37, 52–62. <https://doi.org/10.1016/j.esp.2014.06.005>
- Mur-Dueñas, Pilar. 2007. ‘I/we focus ony’: A cross-cultural analysis of self-mentions in business management research articles. *Journal of English for Academic Purposes* 6, 143–162. <http://dx.doi.org/10.1016/j.jeap.2007.05.002>
- Mur-Dueñas, Pilar. 2010. Attitude markers in business management research articles: A cross-cultural corpus-driven approach. *International Journal of Applied Linguistics* 20 (1), 50–72. <https://doi.org/10.1111/j.1473-4192.2009.00228.x>
- 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>
- Mur-Dueñas, Pilar. 2021a. There may be differences: Analysing the use of hedges in English and Spanish research articles. *Lingua* 260, 103–131. <https://doi.org/10.1016/j.lingua.2021.103131>
- Mur-Dueñas, Pilar. 2021b. Engagement markers in research project websites: Promoting interactivity and dialogicity. *Poznań Studies in Contemporary Linguistics* 57 (4), 655–676. <https://doi.org/10.1515/psicl-2021-0023>
- Park, Hae In & Sinae Lee. 2022. Interactional Metadiscourse in “English Teaching” Articles: A Diachronic Perspective (1980–2021). *English Teaching* 77 (2), 3–23. <http://dx.doi.org/10.15858/engtea.77.2.202206.3>

- Pisanski Peterlin, Agnes. 2005. Text-organising metatext in research articles: an English – Slovene contrastive analysis. *English for Specific Purposes* 24, 307–319. <https://doi.org/10.1016/j.esp.2004.11.001>
- Rahimpour, Sepideh & Esmail Faghieh. 2009. Contrastive rhetoric of English and Persian written texts: Metadiscourse in applied linguistics research articles. *Rice Working Papers in Linguistics* 1, 92–107.
- Ruskan, Anna, Helen Hint, Djuddah Arthur Joost Leijen & Jolanta Šinkūnienė. 2023. Lithuanian academic discourse revisited: Features and patterns of scientific communication. *Open Linguistics* 9/1. <https://doi.org/10.1515/opli-2022-0231>
- Skorzynska, Hanna & María Luisa Carrió-Pastor. 2021. A cross-disciplinary study of verb boosters in research articles from Engineering, Medicine and Linguistics: Frequency and co-text variations. *Revista Signos. Estudios de Lingüística* 54 (106), 575–599. <http://dx.doi.org/10.4067/S0718-09342021000200575>
- Sultan, Abbas HJ. 2011. A contrastive study of metadiscourse in English and Arabic linguistics research articles. *Acta Linguistica* 5 (1), 28–41.
- Swales, John M. 1990. *Genre analysis: English in academic and research settings*. Cambridge, UK: Cambridge University Press.
- Šinkūnienė, Jolanta. 2011. *Autoriaus pozicijos švelninimas rašytiniame moksliniame diskurse: gretinamasis tyrimas*. [Hedging in written academic discourse: A cross-linguistic and cross-disciplinary study]. Humanitarinių mokslų daktaro disertacija [PhD dissertation]. Vilnius: Vilniaus universitetas.
- Šinkūnienė, Jolanta. 2012. Adverbials as hedging devices in Lithuanian academic discourse: A cross-disciplinary study. *Multiple perspectives in linguistic research on Baltic languages*. Aurelija Usonienė, Nicole Nau & Ineta Dabašinskienė, eds. Newcastle upon Tyne: Cambridge Scholars Publishing. 137–167.
- Šinkūnienė, Jolanta. 2014. *Lietuviškojo humanitarinių ir socialinių mokslų diskurso ypatybės*. [Features of Lithuanian discourse in the humanities and social sciences]. Vilnius: Vilniaus universiteto leidykla.
- Šinkūnienė, Jolanta. 2018. The power of English: *I* and *we* in Lithuanian, Lithuanian English and British English research writing. *Intercultural Perspectives on Research Writing*. Pilar Mur Dueñas & Jolanta Šinkūnienė, eds. Amsterdam, Philadelphia: John Benjamins. 59–79.
- Šinkūnienė, Jolanta. 2019. Reformulation markers in academic discourse: A cross-linguistic and cross-disciplinary study. *Jezikoslovlje* 20 (3), 531–553. <http://dx.doi.org/10.29162/jez.2019.19>
- Šinkūnienė, Jolanta. 2020. Konstrukcijos su nederinamuoju neveikiamosios rūšies dalyviu *galima*: diskurso žymiklio link. *Darbai ir Dienos* 74, 77–96.
- Šliogerienė, Jolita, Giedrė Valūnaitė-Oleškevičienė & Vilma Asijavičiūtė. 2015. Discourse relational devices of contrast in Lithuanian and English. *Santalka: filologija, edukologija* 23 (2), 92–100. <http://dx.doi.org/10.3846/cpe.2015.273>

- Thompson, Geoff & Puleng Thetela. 1995. The sound of one hand clapping: the management of interaction in written discourse. *TEXT* 15 (1), 103–27.
- Tse, Polly & Ken Hylad. 2006. ‘So what is the problem this book addresses?’: Interactions in academic book reviews. *Text and Talk* 26, 767–790.
- Usonienė, Aurelija. 2013. On the morphosyntactic status of complement-taking predicate clauses in Lithuanian. *Acta Linguistica Hafniensia: International Journal of Linguistics* 45 (1), 73–99. <https://doi.org/10.1080/03740463.2014.892662>
- Utkā, Andrius. 2009. *Dažninis rašytinės lietuvių kalbos žodynas: 1 milijono žodžių morfologiškai anotuoto tekstyno pagrindu*. [Frequency dictionary of the written Lithuanian language: based on the morphologically annotated corpus of 1 million words]. Kaunas: Vytauto Didžiojo universiteto leidykla.
- Vande Kopple, William J. 1985. Some exploratory discourse on metadiscourse. *College Composition and Communication* 36, 63–94.
- Vázquez, Ignacio & Diana Giner. 2009. Writing with conviction: The use of boosters in modelling persuasion in academic discourses. *Revista Alicantina de Estudios Ingleses* 22, 219–237. <http://dx.doi.org/10.14198/raei.2009.22.14>