TURJE
www.turje.org

2024
Volume: 13 Issue: 3

Turkish Journal of Education

https://doi.org/10.19128/turje.1383179
Research Article

Received 30.10.2023
Accepted 28.05.2024

A corpus-based analysis of critical thinking through interactional metadiscourse in pre-service EFL teachers' writing

Selahattin Yılmaz
Yıldız Technical University, School of Foreign Languages, İstanbul, Türkiye, yilmazs@yildiz.edu.tr
Ferda İlerten
Yıldız Technical University, Department of Foreign Languages Education, İstanbul, Türkiye, filerten@yildiz.edu.tr

ABSTRACT Critical writing seeks to enhance university students' ability to think causally and reason effectively, and this improvement should be evident in their language use in the assignments. An example of such language is interactional metadiscourse, the expression of attitudes and opinions in line with the intended audience. In pursuit of these objectives, this study investigated the textual characteristics of critical thinking by examining interactional metadiscourse markers (MDMs) in the critical response papers authored by English Language Teaching (ELT) undergraduate students throughout a semester at a Turkish state university. The findings revealed shifts in the use of interactional MDMs by the end of the semester. While markers for engagement, hedging, and boosting remained prevalent across various tasks, the utilization of self-mentions and attitude markers declined, indicating a transition from the students' sharing personal opinions and experiences to relying on evidence from research in academic texts to support their arguments. Additionally, the study highlighted the impact of topic selection on how students incorporated metadiscourse markers into their response papers.

Keywords: *Critical thinking, English for academic purposes, Metadiscourse, Pre-service teacher education, Second language writing*

İngilizce öğretmen adaylarının yazımlarında etkileşimsel üstsöylem yoluyla eleştirel düşünmenin derlem tabanlı analizi

ÖZ Eleştirel yazma, üniversite öğrencilerinin nedensel düşünme ve akıl yürütme becerilerinin geliştirilmesini hedeflemektedir ve bu gelişimin öğrencilerin ödevlerinde kullandıkları dile yansıması beklenmektedir. Bu dil kullanımının bir örneği, hedeflenen kitleye uyumlu tutum ve görüşlerin ifade edildiği etkileşimsel üstsöylemdir. Bu hedefler doğrultusunda, bu çalışma, Türkiye'deki bir devlet üniversitesinde bir akademik yarı yıl boyunca İngilizce öğretmen adayları tarafından yazılan eleştirel değerlendirme raporlarındaki etkileşimsel üstsöylem belirteçlerini (EÜB'ler) inceleyerek eleştirel düşünmenin metinsel özelliklerini araştırmıştır. Bulgular, dönem sonunda söylem işaretleyicilerinin kullanımında anlamlı farklılıklar olduğunu ortaya koymuştur. Katılım, kaçınma ve güçlendirme belirteçleri bazı raporlarda yaygın olarak kullanılırken, bazı raporlarda ise kendinden bahsetme ve tutum belirteçlerinin kullanımı azalmıştır. Elde edilen bu sonuç, öğrencilerin kişisel görüşlerini ve deneyimlerini paylaşmaktan, argümanlarını desteklemek için akademik metinlerdeki araştırma bulgularına güvenmeye geçiş yaptıklarını göstermektedir. Ek olarak çalışma, öğrencilerin eleştirel değerlendirme raporlarına EÜB'leri nasıl dahil ettikleri hususunda konu seçiminin etkisini vurgulamıştır.

Anahtar Sözcükler: *Akademik İngilizce, Eleştirel düşünme, Hizmet öncesi öğretmen eğitimi, İkinci dilde yazma, Üstsöylem*

Citation: Yılmaz, S., & İlerten, F. (2024). A corpus-based analysis of critical thinking through interactional metadiscourse in pre-service EFL teachers' writing. *Turkish Journal of Education, 13*(3), 239-265. https://doi.org/10.19128/turje.1383179

INTRODUCTION

Conveying field-specific knowledge in linguistically and structurally appropriate ways is one of the challenges of college-level writing (Hyland, 2013). The task gets more complicated in second language (L2) academic writing, where university students deal with the cognitive demands of their discipline and the task of writing in a foreign language (Breeze, 2012). Once they begin using explicit signals of communication, the message in the text can be evaluated by the reader more efficiently. Thus, metadiscourse is one of the crucial constituents of writing instruction for academic purposes as it facilitates the transmission and comprehension of messages (Hyland, 2005a).

Metadiscourse refers to the textual properties used to organize discourse and align a writer's stance with the content and intended audience (Hyland, 2005a, p. 14). Markers in texts reveal the author's personality, beliefs, and attitudes towards the content, making them crucial communication tools beyond information transmission (Trillo, 2002). Among the different taxonomies of metadiscourse markers (henceforth, MDMs), a commonly used one was proposed by Hyland and Tse (2004) following Thompson and Thetela's (1995) two-way model of interaction in writing. According to the model, interactive MDMs, including transitions (e.g., in addition, but), frame markers (e.g., finally, to conclude), endophoric markers (e.g., noted above, see Fig), evidentials (e.g., according to X, Z states), and code glosses (e.g, namely, such as), help the reader navigate through the text, and interactional MDMs, including hedges (e.g., might, perhaps), boosters (e.g., in fact, definitely), attitude markers (e.g., unfortunately, I agree), engagement markers (e.g., consider, note that), and self-mentions (e.g., I, we), are used to adjust the reader's involvement in the text (Hyland & Tse, 2004, p. 169). While the former deals mostly with textual organization, the latter is key to understanding the writer's way of interacting with the reader, evaluating propositions, and expressing their stance (Hyland, 2005b). One of the seminal works using Hyland and Tse's (2004) model is Hyland's (2005a) monograph, in which the researcher not only comprehensively revisited the earlier model but also explored the use of metadiscourse features across rhetorical practices, genres, cultures, and discourse communities, and provided a comprehensive list of 300 potential MDMs. The series of studies in the monograph showcases that metadiscourse could play a crucial role in understanding interactions, particularly in written academic and professional communication, thus having important implications for writing pedagogy.

Another significant variable in effective college-level writing in a foreign language is the reflection of critical thinking skills on writing assignments (McKinley, 2013). Although the term critical thinking has been regarded as an essential skill to acquire in higher education, the perception and definition of it vary across disciplines. Moore (2013) demonstrates the aspects various academic disciplines such as philosophy, history, and cultural studies focus on in their interpretation of critical thinking. Yet, in the study, common themes emerged as judgment, a skeptical view of knowledge, simple originality, sensitive reading of a text, rationality, adopting an ethical stance, and self-reflectivity. Similarly, Ennis (1987) defines critical thinking as "reasonable, reflective thinking that is focused on deciding what to believe or do" (p. 10). It involves a chain of terms such as problem-solving, metacognition, decision-making, rational thinking, and reasoning (Ennis, 1992). The present study adopts Bruce's (2018) definition of critical thinking as "an evaluative judgment made within any field of human activity about some aspect, object, or behavior of that field" (p. 4). Bruce also adds that, following Swales and Feak (2012), these judgments should conform to the established standards of that field.

Studies have sought to identify the important characteristics of critical thinking, and how they can be traced in writing. For instance, Facione (2011) lists several skills key to critical thinking, involving interpretation of information, analysis, inference, evaluation, explanation, and self-regulation. Similarly, Nosich (2021) illustrates how qualities of good reasoning, such as clarity, accuracy, and relevance, can manifest themselves in critical writing. The exploration of how critical thinking translates into writing as a cognitive ability (Liu & Stapleton, 2018) has been investigated in several research studies as well (Carroll, 2007; Condon & Kelly-Riley, 2004; Woodward-Kron, 2002). Carroll (2007) investigated college students' term papers at the beginning and end of the semester to observe linguistic and cognitive development. Similarly, Woodward-Kron (2002) evidenced that descriptive pieces of essays (i.e.,

naming, taxonomizing, etc.) were preliminary steps to writing and considering the subject matter in a critical manner. It could be deduced from the studies that the acquisition of critical thinking and writing skills follows an incremental process. Further, previous studies point to a reciprocal relationship between critical thinking and writing skills. As put forward by Carroll (2007), college-level writing boosts students' field-specific knowledge and critical thinking skills. In return, critical thinking skills can be reflected in writing through the strong use of arguments and MDMs. Bruce (2018) suggests that interactional metadiscourse in particular is an analytical instrument that can be used to observe the text-level characteristics of critical thinking, such as how a writer persuades the reader and expresses attitudes and ideas efficiently. For instance, writers can convey their opinions or reduce the power of their claims by using hedges (Liu & Stapleton, 2018). According to Liu and Stapleton (2018), although the use of hedges reduces the volume of the writer, it can be regarded as a higher order thinking skill as the readers' interpretation and perception are also considered.

Previous Studies on the Use of Metadiscourse in L2 Academic Writing

In recent years, there has been a proliferation of metadiscourse research in L2 academic writing, including areas such as expert writing (e.g., Dahl, 2004; Gillaerts & Van de Velde, 2010); graduate and undergraduate-level writing (e.g., Akbaş, 2014; Bayyurt, 2010; Çandarlı et al., 2015; Ho & Li, 2018; Hyland, 2004; Lee & Deakin, 2016); learner writing (e.g., Ädel, 2006; Qin & Uccelli, 2019; Yoon, 2021); longitudinal studies (e.g., Beyazyildirim & Ercan, 2023; Crosthwaite & Jiang, 2017; Gürsoy, 2023; Martin-Laguna, 2023; Ruan, 2019); diachronic studies examining the change in metadiscourse over time (e.g., Hyland & Jiang, 2018). Several studies have also investigated the features of critical thinking and writing (e.g., Bruce, 2016; Carroll, 2007; Lancaster, 2016; Liu & Stapleton, 2018; Woordward-Kron, 2002). In line with the scope of the research, studies that reflect the developmental pattern, selection of writing prompts/materials, and characteristics of the discourse community are covered in this section.

Among the longitudinal studies in L2 writing development, Crosthwaite and Jiang (2017) examined the role of EAP (English for Academic Purposes) instruction in the semester-long development of stance devices, i.e., hedges, boosters, self-mentions, and attitude markers of Hyland (2005a), in undergraduate freshman writing. They found that the students made use of fewer stance features, signaling a more cautious and impersonal academic style. In line with academic conventions, they were also able to use certain devices, such as boosters, in a more controlled way in cases where stance expression was necessary. Another longitudinal metadiscourse study is Ruan (2019), which compared the use of MDMs, according to Hyland's (2005a) model, in undergraduate essays by three groups over time: Chinese English Medium Instruction (EMI) students, Chinese English majors, and first language (L1) English writers. The study reported that, compared to the other two groups, the Chinese English majors used MDMs twice as often. The differences were most pronounced in self-mentions, engagement markers, and boosters. As for the longitudinal changes, while the L1 English writers' metadiscourse use remained stable, a decrease in the frequency of engagement markers and self-mentions in EFL and an increase in the frequency of hedges and boosters in EAP were reported. More recently, Martin-Laguna (2023) examined the use of attitude markers and hedges by intermediate Catalan-Spanish bilingual learners of third language (L3) English in three opinion writing tasks over the course of an academic year. The findings indicated that the use of hedges increased and attitude markers decreased significantly from the first to the third writing task. Furthermore, the researchers also highlighted that the increase in hedges was steady, while that in attitude markers exhibited fluctuation.

A number of studies focused on the differences in MDM use between high-graded and low-graded writing. Employing Hyland's (2005a) taxonomy, Lee and Deakin (2016) compared the use of interactional MDMs in successful L1 and L2, as well as less successful L2 argumentative essays by undergraduate students. They found that both L1 and L2 successful essays made more frequent use of hedging devices and that L2 writers in general tend to refrain from expressing their identity and stance, which is evident in the infrequent use of particular self-mentions. In another study, Ho and Li (2018) investigated metadiscourse in 181 timed argumentative essays written by Chinese undergraduate

English writers, finding that engagement markers and hedges were the top two commonly used interactional MDMs. Essays with higher ratings exhibited greater diversity and a higher frequency of hedges and attitude markers compared to essays with lower ratings, which were also found to have a much higher frequency of engagement markers.

Other studies delved into the analysis of how MDMs are used across different registers, L1s and proficiency levels. For instance, in Qin and Uccelli's (2019) investigation of the use of MDMs in personal emails and academic essays of EFL learners based on Hyland's metadiscourse model (2005a), the distributional map of MDMs in the EFL learner corpus, the variability of MDMs across the registers, and the relationship between registers and writing quality showed that code glosses were commonly used in academic writing, yet there was an increasing number of boosters, hedges, engagement markers, and self-mentioning in informal writing. Hedges were more frequent in graduate students' academic than colloquial writing, presumably because they welcomed alternative opinions and were more associated with the academic discourse than undergraduate students. Yoon (2021), on the other hand, investigated the use of metadiscourse categories (Hyland, 2005a) among different L1 backgrounds (Korean, Japanese, and Chinese EFL students), topics, and L2 proficiency levels, and through a post hoc analysis, compared their essays with native speaker writing. The results showed that topic and L1 background played a significant role in the use of interactional MDMs. Korean students used fewer hedges and boosters, Japanese students used more stance and engagement markers, and Chinese students' essays included more reader pronouns and fewer self-mentions. However, the use of MDMs was not significantly influenced by L2 proficiency. When their essays were compared with native speaker writing, the underuse of hedges and relatively more frequent use of reader pronouns were observed in the EFL group.

A growing body of literature has also examined essays written by Turkish students in L2 English and L1 Turkish (e.g., Algı, 2012; Bayyurt, 2010; Beyazyildirim & Ercan, 2023; Çandarlı et al., 2015; Gürsoy, 2023; Yüksel & Kavanoz, 2018). These studies focused on the norms of the discourse community and their reflections on student writing. For instance, Bayyurt's (2010) analysis of undergraduate essays and Algı's (2012) analysis of preparatory-level students' paragraphs showed that Turkish students wrote in a cautious manner by including more frequent use of hedges than boosters, which were generally more frequent in L1 Turkish essays. Additionally, Çandarlı et al. (2015) investigated the use of authorial presence markers in argumentative essays by Turkish and American students, followed by interviews with Turkish students. They reported a lower frequency of first-person pronouns and a higher frequency of attitude markers in both English and Turkish essays by Turkish writers than those by American students. The researchers concluded that Turkish essays included more authorial presence markers than English essays written either by Turkish or American students. Alternatively, Yüksel and Kavanoz (2018) examined the use of MDMs in essays by non-native (Turkish L1) and native undergraduate writers, as well as published academic texts by native writers. Their findings showed that interactive MDMs were used more commonly than interactional MDMs. In addition, while the most common interactional MDM category was engagement markers in all three corpora, overuse of self-mentions and underuse of hedges in L1 Turkish writers' texts were reported in comparison to native writing. In their small-scale longitudinal study on argumentative writing by preparatory-level English language learners, Beyazyildirim and Ercan (2023) reported that the most frequent interactional MDM types were engagement markers followed by hedges in the analyzed texts and that the overall frequency of the interactional MDMs decreased as the students' proficiency increased. One noticeable exception was the engagement marker we, which the authors highlighted as a sign of increased reader involvement in student texts. Similarly adopting a longitudinal design, Gürsoy (2023) focused on the use of hedges in nine different argumentative essays written by undergraduate Turkish writers of English in a semester, alongside explicit instruction on hedges. The researcher found that modal auxiliaries such as should remained the most frequent hedging devices, and the overall frequency of hedges increased throughout the semester.

Previous Studies on Critical Thinking through the Use of Metadiscourse

While the literature extensively scrutinizes critical thinking at the college level, its examination through a metadiscourse lens remains relatively unexplored. In a limited number of studies, the reflection of critical thinking skills on student essays was observed in themes such as the development of linguistic and cognitive skills in writing, the development of stance expression, and the effects of prompts on student writing. This section aims to provide both an overview of fundamental perspectives on critical thinking and a discussion of pertinent studies that contribute to this field.

Firstly, Carroll (2007) assessed the impact of a critical thinking course on students' linguistic and cognitive development over a semester. Results showed more linguistic changes than cognitive ones, with students using fewer words for insight, discrepancy, and tentativeness and more for inhibition and causal thinking. Personal pronouns decreased, and certainty increased. Overall, the study demonstrated that students began with a subjective view and advanced to a more constructive perspective at the end. More recently, as part of their genre approach to critical thinking, Bruce (2016) qualitatively examined metadiscourse features in L1 literature and sociology essays, focusing on hedges, attitude markers, boosters, and self-mentions. The study found that hedges and attitude markers constituted the majority of MDMs in both disciplines, while the former was found to be much less frequent in sociology essays. Bruce concluded that both categories of interactional MDMs are key to the effective expression of critical stance in essays. Shifting their focus to stance expression in undergraduate writing, Lancaster (2016) analyzed argumentative essays from economics and political theory courses. High-graded essays used more hedges, boosters, attitude markers, and disclaim markers, while low-graded ones featured more self-mentions. According to the author, the effective use of these markers in high-graded essays allowed students to position themselves as novice writers, question ideas, detach from the text, and minimize personal involvement, reflecting clear and effective critical thinking. Liu and Stapleton (2018) investigated the effect of prompt differences on the writing performance and critical thinking of Chinese undergraduate students through MDM use. They found that hedges and attitude markers were significantly more frequent among student texts written in response to the discipline-specific problem-solving prompt in comparison to a traditional exam prompt. As can be seen, although limited in number, these studies were able to describe certain important qualities of critical thinking through undergraduate writing.

To summarize, the aforementioned studies show that the use of interactional MDMs by undergraduate students is a complex phenomenon influenced by many factors, such as L1 background, instruction, writing prompt, context, and discipline. Shedding light on different dimensions of interactions in text, interactional MDMs can also offer valuable insights into the critical thinking processes of L2 students, as evidenced in their writing. Following these investigations, the present study aims to analyze the progression of critical thinking skills in student essays while also identifying the key aspects that influence their writing.

The Current Study

This study investigates the textual characteristics of critical thinking by examining interactional MDMs in Hyland (2005a) in critical response papers authored by pre-service EFL teachers throughout a semester at a Turkish public university. In line with the aims of the study, the following research questions were addressed:

(1) What are the commonly used interactional MDMs as listed in Hyland (2005a), in the critical response papers by pre-service EFL teachers?

(2) How does the use of interactional MDMs differ across critical response writing tasks?

METHODOLOGY

Research Context and Writing Tasks

The study involved 22 participants, five male and 17 female students aged 19–20 years, in an English Language Teaching (ELT) undergraduate program at a state university in Istanbul, Türkiye. All the participants were L1 Turkish speakers who were non-native users of English at upper-intermediate to advanced proficiency levels. The data comprise the writing assignments of the "Critical Reading and Writing (CRW)" course. The course is offered to sophomore-level students in the four-year program. It is a compulsory course in ELT programs in Türkiye and is taught two hours per week in the fall semester. According to the latest regulation of English Language Teacher Education Programs (Council of Higher Education, 2018), the aim of the CRW course is:

> *To be able to analyze, summarize and/or report current studies selected from the field of English language education; examining studies in their context and localize knowledge; comparing and synthesizing texts advocating different views on the subject and producing their own original texts (p. 8).*

As the course was taught during the Covid-19 pandemic, all sessions were held online. Throughout the semester, texts on nine current issues in ELT were examined weekly. A different theme was covered each week throughout the course. The students wrote a response essay based on the readings and class discussions around these themes (see Appendix 1 for further details of the weekly discussions).

The sessions were followed by the critical response writing process and their submission in the two weeks following the course. The students were asked to respond to the articles read and discussed in class every week, by first summarizing the texts and then critically reflecting on them. Except for the last essay, the instructor provided the reading materials and asked participants to write critical response papers based on the topics given. For the last essay, which was graded as the final paper, brief information was given to the participants about the topic, and they were asked to find two articles and write their critical essays on them. Their essays were assessed based on the robustness of the introduction and conclusion, the structure of the body paragraph, the effective conveyance of ideas, organization, style, and mechanics.

The critical response papers were collected from the participants after the necessary approval from [NAME OF UNIVERSITY] Human Research Ethics Committee (Date: 02.06.2022, Project ID: 000527, Ethics Committee Meeting Number: 2022.06) was obtained. Of the 81 students enrolled in the CRW course, 22 students who submitted all nine papers were included in the study. Thus, the small corpus of critical response papers, hereafter referred to as the CRPC, comprises a total of 198 documents. 22 CRPs written for each theme were used as a subcorpus. Table 1 presents a thematic and numerical overview of the CRPCs. As can be seen, the subcorpora are of generally similar length, but the papers in the subcorpora of the fourth, sixth, and eighth tasks are relatively shorter. In addition, the type-token ratios for the subcorpora are somewhat similar, pointing to a comparable level of lexical diversity.

In contrast to large corpora that capture general language use, small specialized corpora such as CRPC have certain important advantages. First, highly specialized research questions as in the case of this study generally necessitate small, tailor-made corpora in order to be addressed more accurately (Ross, 2018). Moreover, a small and specialized corpus could enable "a much closer link between the corpus and the contexts in which the texts in the corpus were produced" (Koester, 2010, p. 67). For instance, being the instructor of the course from which the data were collected, the second author provided valuable insights into the interpretation of the findings. In this way, factors such as the participants' profile and background knowledge as well as the role of topic selection were also taken into account. Besides, limiting the study to texts written in response to nine consecutive tasks by the same students in one semester allowed a certain degree of homogeneity in the data and more focused results. Thus, despite its relatively small size, CRPC provides a situated understanding of how critical thinking is textually

manifested through interactional metadiscourse, which can inform larger-scale corpus studies in the future.

Table 1.
Overview of the CRPC

Subcorpus	Theme of CRPs	Average length	Word count	TTR(%)
T1	Use of technology in ELT	990.5	21791	40.08
T2	Mindfulness and social-emotional learning	865.55	19042	40.64
T3	What do successful students and teachers do?	909.32	20005	38.93
T4	Translanguaging	684	15048	42.75
T5	Do-it-yourself (DIY) in ELT	896.59	19725	37.54
T6	Online ESL/EFL classes	737.73	16230	41.96
T7	Advocacy in language classes	904	19888	40.41
T8	Teaching proficiency through reading and storytelling	767.41	16883	41.7
T9	Raising bilingual children	977.18	21498	37.47
Total		859.14	170110	40

Data Analysis

Firstly, over 20 percent of the data, that is, papers by five participants, was manually annotated by both researchers to get a deeper understanding of the language used by the participants in their papers. The degree of agreement between the researchers was assessed via the percentage agreement method, that is, the number of times both rates agreed divided by the total number of items rated. Upon comparing the annotated parts, an agreement rate of 86.30 % was reached, which is above the 80% threshold recommended by Fraenkel et al. (2023). The annotated parts included markers such as modal verbs, personal pronouns, adverbs, and adjectives, signaling the participants' stance towards the papers, degree of certainty and commitment to their claims, and interactions with the reader. After reviewing the relevant literature, the researchers noticed a remarkable overlap between the manually annotated markers and those that have long been the focus of metadiscourse research. Therefore, to comprehensively analyze these patterns, Hyland's (2005a) model of interactional metadiscourse was adopted for the current study (see Table 2). Since this study included the complete list of MDMs in the model, this list was employed to automatically extract interactional MDMs.

Table 2.
A Model of Interactional Metadiscourse (Adapted from Hyland, 2005a, p. 49)

Category	Function	Examples
Hedges	withhold commitment and open dialogue	*might, perhaps, possible, about*
Boosters	emphasize certainty or close dialogue	*in fact, definitely, it is clear that*
Attitude markers	express writer's attitude toward proposition	*unfortunately, I agree, surprisingly*
Self mentions	explicit reference to author(s)	*I, we, my, me, our*
Engagement markers	explicitly build relationship with reader	*consider, note, you can see that*

First, the participants' personal information was removed from the texts, and all the spelling mistakes were corrected using Microsoft Word. As participants were highly proficient L2 users, no added error tagging or annotation was made. Next, EncodeAnt 1.2.1 (Anthony, 2017) was used to convert the documents into UTF-8 encoded text (.txt) files for corpus analysis. Then the markers in Hyland's (2005a) full list of interactional MDMs were searched by using the "advanced search" feature of the concordance tool in AntConc 3.5.9 (Anthony, 2020). To use this tool, the full list of MDMs, along with all possible variations of markers such as verbs (e.g., all inflectional forms of the verb argue in hedges) were imported into the program as a text document. The researchers then adopted a vertical reading methodology and carefully examined the analysis outputs by reading the concordance lines. However, it is of great importance that such vertical reading of data be supported by horizontal reading (Aijmer & Rühleman, 2015). Therefore, the researchers also did a horizontal reading in cases where further context is needed to better understand the MDM's functions.

While most interactional MDMs in Hyland's (2005a) list were included in the study, the list was slightly

adapted by making several changes in the way certain markers were categorized. For instance, all CRPs were written individually, so all first-person plural pronouns, i.e., we, us, our, were coded as engagement markers used to invite fellow teachers and students into the discussion in essays. In addition, should was categorized under both hedges and engagement markers in Hyland's (2005a) model. However, since no instance of should as a hedging device was found in our data, it was only listed as an engagement marker.

The extracted frequencies of MDMs were normalized per 10,000 words across the nine subcorpora. Similar to Brezina (2018), normalized frequencies of MDMs were used for statistical tests due to the carrying length of CRPs. Jamovi 2.2 (The Jamovi Project, 2021) was used to run repeated measures Analysis of Variances (ANOVAs) and following post hoc tests for pairwise comparisons. Because of the small sample size (N=22) and non-normality of distribution detected for the independent variables, non-parametric tests were used. Therefore, Friedman's (1937) non-parametric repeated measures ANOVAs were used to investigate the differences in the frequency of metadiscourse use across writing assignments. A different one-way ANOVA was run for each index of interactional metadiscourse. Next, pairwise comparisons were made via a series of Durbin-Conover (Conover, 1999) post hoc tests, also available in Jamovi through Pohlert's (2018) PMCMR R Package. Next, Bonferroni adjustments were made for the multiple comparison tests ($\alpha = .05/36 = .001$). Because non-parametric tests were used, median (M) and median absolute deviation (AD) scores were reported as descriptive statistics.

FINDINGS

The analysis included the interpretation of both overall frequencies of interactional MDMs, and five individual categories, namely, attitude markers, engagement markers, self-mention, hedges, and boosters.

Overall Frequencies

As seen in Table 3, engagement markers have the highest token frequency in CRPC, which is followed by hedges, boosters, and self-mentions, respectively. Lastly, the least frequent markers are the attitude markers. As for the distinct types of markers used, attitude markers have the highest type frequency, indicating a great number of different attitude markers with low frequencies. Having the least type frequency with only four distinct first-person pronouns, self-mentions are the least lexically diverse group of interactional MDMs (see Appendix 2 for the 10 Most Frequent Interactional MDMs).

Table 3.
Overall Frequencies, Type-Token Ratios (TTRs), and the 10 Most Frequent MDMs

Category	Type	Token (raw)	Token (normed)	TTR (%)	The most frequent 10 types
Attitude markers	35	938	55.14	3.73	*important, even, agree, appropriate, essential, interesting, prefer, disagree, understandable, unfortunately*
Boosters	33	1662	97.7	1.99	*think, find, know, believe, must, show, always, really, clear, clearly*
Engagement markers	71	4892	287.58	1.45	*should, we, use, do not, our, you, take, must, ?, develop*
Hedges	60	2132	125.33	2.81	*may, would, could, feel, might, possible, in my opinion, claim, rather, suggest*
Self-mentions	4	1495	87.88	0.27	*I, my, me, mine*

When comparing these frequencies with previous studies, several intriguing observations emerge. The finding that almost half of the identified MDMs were engagement markers have the highest type frequency echoes Tasso's (2020) analysis of L2 English essays by Spanish learners. Furthermore, the ranking of token frequencies aligns completely with Beyazyildirim and Ercan's (2023) findings based on Turkish preparatory class students' argumentative writing. In addition, the prevalence of engagement

markers as the most frequent type, and the relatively infrequent occurrence of attitude markers mirror the findings of Yüksel and Kavanoz (2018). However, the current study reports a much lower frequency of self-mentions, similar to the results of Çandarlı et al. (2015). Furthermore, our findings corroborate Bayyurt (2010) and Algı (2012) in that Turkish writers of English tend to use hedges more frequently than boosters. Lastly, the normed frequency of hedges, the second most frequent interactional MDM group in the current study, is rather close to that in successful L2 essays and L1 essays, as was reported in Lee and Deakin (2016).

Findings Across Tasks

When the normed frequencies of the interactional MDMs in question were examined, it can be seen that the frequencies of the markers fluctuate considerably across tasks (see Table 4). The overall frequency of the MDMs was found to be higher in the second, third, and fifth tasks, whereas the frequencies decreased in the eighth and ninth tasks, which is similar to Beyazyildirim and Ercan's (2023) finding. A similar pattern can be observed in the case of the attitude markers, but the developmental patterns are rather complex for the other categories. Nevertheless, despite the fluctuations, the most steady decline is seen in the case of self-mentions. Another interesting observation is the sudden frequency changes in certain tasks, such as the decrease in attitude markers in the fourth task and hedges in the eighth task, as well as the increase in engagement markers in the second task. In addition, the highest normed frequencies of attitude markers and boosters were found in the third task.

Table 4.
Normed Frequencies of Interactional Metadiscourse Markers (MDMs) Across Tasks

Interactional MDMs	T1	T2	T3	T4	T5	T6	T7	T8	T9
Attitude Markers	59.66	56.72	76.48	34.56	68.44	57.3	47.77	49.16	41.4
Boosters	90.86	104.51	114.97	103	97.85	92.42	94.53	80.55	99.08
Engagement Markers	261.12	384.94	309.92	287.75	325.48	311.77	271.02	209.68	230.72
Hedges	137.21	108.71	114.97	156.17	153.61	118.92	127.72	88.85	121.41
Self-mentions	139.51	84.02	98.48	96.36	80.1	80.71	68.38	74.63	64.19
SUM	688.36	738.9	714.82	677.84	725.48	661.12	609.42	502.87	556.8

Tables 5 and 6 show the longitudinal distribution of interactional MDMs across the nine writing tasks in the form of descriptive (median and absolute deviations) and inferential (repeated measures ANOVA) statistical results. ANOVA tests produced significant differences across tasks in all five interactional MDM categories *($p < .05$)*. However, in line with the fluctuating normed frequencies of the markers across tasks, Bonferroni-adjusted significant pairwise differences were found only between some of the tasks *($p <. 001$)*.

Table 5.
Median and Median Absolute Deviation (AD) Values for Interactional Metadiscourse Markers (MDMs)

	Attitude Markers		Boosters		Engagement Markers		Hedges		Self-mention	
Writing Tasks	M	AD	M	AD	M	AD	M	AD	M	AD
T1	54.45	17.19	94.37	34.16	213.15	49.23	130.85	29.54	146.51	35.4
T2	60.72	13.58	108.5	23.45	335.83	137.24	101.21	26.66	87.16	55.18
T3	73.13	26.66	111.89	37.97	328.91	100.61	112.37	35.38	98.15	58.9
T4	29.3	12.45	97.22	25.94	270.44	44.71	139.62	47.36	65.46	48.56
T5	64.17	20.7	93.66	29.34	321.3	64.98	127.63	44.06	47.21	36.85
T6	54.39	17.9	85.06	33.65	282.92	43.14	107.24	39.83	62.92	48.08
T7	46.06	17.85	86.6	37.39	275.56	81.21	119.44	39.35	54.39	24.7
T8	46.83	24.51	68.42	20.79	215.35	63.34	90.94	35.99	39.54	36.61
T9	36.79	13.8	103.45	18.54	222.78	73.17	114.76	19.19	41.98	27.54

Table 6.
Results of Repeated ANOVA Tests for Categories of Interactional Metadiscourse Markers (MDMs)

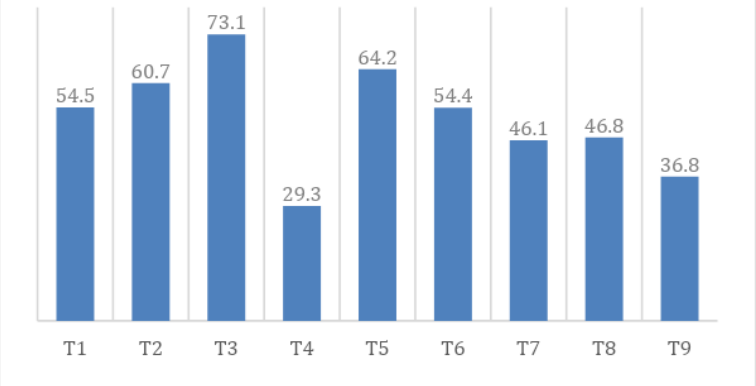
	ANOVA Results			*Pairwise comparisons (significant at .001 level after Bonferroni correction)*
Category	χ^2	DF	P	
Attitude Markers	31.62	8	<.001	T1>T4; T1>T9; T3>T4; T3>T9; T4<T5; T4<T6; T5>T9
Boosters	17.18	8	0.028	T3>T8
Engagement Markers	35.02	8	<.001	T1<T2; T2>T8; T2>T9; T3>T8; T5>T8; T6>T8
Hedges	16.58	8	0.035	T5>T8
Self-mentions	31.74	8	<.001	T1>T6; T1>T7; T1>T8; T1>T9

Note. Only significant results for the pairwise comparisons were reported.

Attitude Markers

A significant difference was found across tasks regarding the frequency of attitude markers. Having the lowest median frequencies, the fourth and ninth tasks differed significantly from most other tasks (Fig. 1). The frequency of attitude markers was found to be significantly smaller than those in the first, third, fifth, and sixth tasks. Similarly, the use of attitude markers was significantly more frequent in the first, third, and fifth tasks than in the ninth task. Similar to Liu and Stapleton's (2018) finding that MDM use changes across writing prompts, it is plausible to suggest that one possible explanation for this difference is the writing topic. Compared to the other more practical topics generally discussed in other courses, participants were not as familiar with translanguaging, and child bilingualism, i.e., the topics for the fourth and ninth tasks, respectively. It is also worth mentioning that, as the second author, who was the instructor of the course, points out, the participants were skeptical of the role of translanguaging in language teaching. Our findings support those of Martin-Laguna (2023) in terms of the fluctuating frequencies as well as the eventual decrease in the frequency of attitude markers. Having similarly reported a progressive decline in the use of attitude markers, Crosthwaite and Jiang (2017) attributed this finding to "an emphasis on justifying claims rather than appealing to emotion" (p. 98).

Figure 1.
Median Frequencies of Attitude Markers Across Tasks (per 10.000 words)



*If they practice it continuously, they can easily observe the positive effects **even** in a short time like 2-3 weeks (T2, S11)*

*I agree with the authors that an effective material should be **appropriate** for the student level and native language. (T5, S35)*

*However, as with any tool, it is **important** to know its proper uses, and when it should not be used. (T6, S61)*

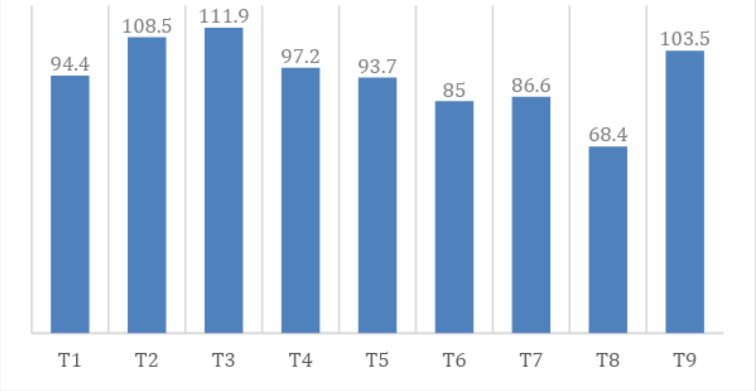
*I **agree** with this view because the importance and support that families attach to education cause the child to better understand the importance of education….(T7, S35)*

*It is also **important** to note that it is essential to design various kinds of activities and tasks in order to provide various learning styles. (T8, S28)*

Boosters

Although a significant difference was also found in the use of boosters across tasks, the only post hoc pairwise difference was reported between the third and eighth tasks. The former has the highest median frequency of boosters, while the latter has the lowest, which could also be linked to the choice of topic (Fig. 2). The discussion of what successful teachers and learners should do in the third task appears to have given the participants the chance to voice their opinions and provide recommendations, adopting a rather strong stance at times. This is in line with increased certainty and boosting claims in critical writing tasks in the literature (e.g., Carroll, 2007; Lancaster, 2016). However, the discussion of "Teaching Proficiency Through Reading and Storytelling (TPRS)" in the eighth task, a teaching approach with which most participants were unlikely to be familiar, boosting their claims, wasn't deemed necessary. In addition, the median frequencies from the third to the eighth task point to a slight and somewhat progressive decrease in the use of boosters, which is similar to Crosthwaite and Jiang's (2017) finding following explicit metadiscourse instruction. Nonetheless, the declining pattern was not visible in the first three and ninth tasks.

Figure 2.
Median Frequencies of Boosters Across Tasks (per 10,000 words)



*I **think** that all these possibilities should be taken into consideration before using technology in education. (T1, S66)*

*Furthermore, the teachers **must** be resilient as the world and knowledge change swiftly. (T3, S26)*

*We **know** that everyone's language learning speed is different from each other. (T4, S8)*

*However, I **found** the article inadequate in terms of the way the subject was dealt with. (T8, S35)*

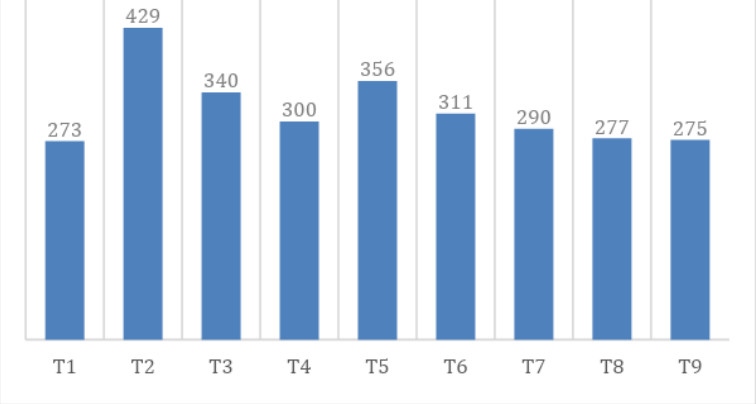
*….and I **believe** that this is not a disadvantage of raising bilingual children (T9, S28)*

Engagement Markers

As for the frequency of engagement markers, pairwise comparisons following the significant differences across tasks pointed to two tasks differing significantly from several others. A significant difference was reported between the second writing task on mindfulness and social-emotional teaching with the highest median frequency of engagement markers of all tasks and the three tasks with the lowest median frequencies, i.e., the first, eighth, and ninth tasks (Fig. 3). In addition, the use of engagement markers in the eighth task was significantly lower than in the second, third, fifth, and sixth tasks. This shows that the participants employed a more engaged style when writing about mindfulness and social-emotional teaching, and they were least engaged in their written discussion of TPRS, which also included the

scarcest use of boosters. This is also evident from the stark difference in the frequencies of "we" and "should", both of which have remarkably higher frequencies in the second task in comparison to the eighth. Both items are generally used together by the students to offer teachers advice by also identifying themselves as one of them. In addition, it is also clear from median frequencies across tasks that the participants made slightly less frequent use of engagement markers toward the end of the semester, which is similar to Ruan's (2019) results from their analysis of essays by Chinese English major students. Nevertheless, engagement markers remained to be the most frequent MDMs across all tasks, which could be interpreted as the student's willingness to engage the reader in their texts (Tasso, 2020).

Figure 3.
Median Frequencies of Engagement Markers Across Tasks (per 10,000 words)



*We **should** pay attention to these elements, if we really want our students to learn effectively and to motivate them in their language learning journey. (T2, S28)*

*….an effective teacher **should** be able to provide the right instructions and help students to develop their knowledge, skills and understandings…. (T3, S7)*

*….so I think **we** can use this technique in education to strengthen the communication of students. (T4, S35)*

*Therefore, I **do not** think that online teaching can be superior to face-to-face education. (T6, S66)*

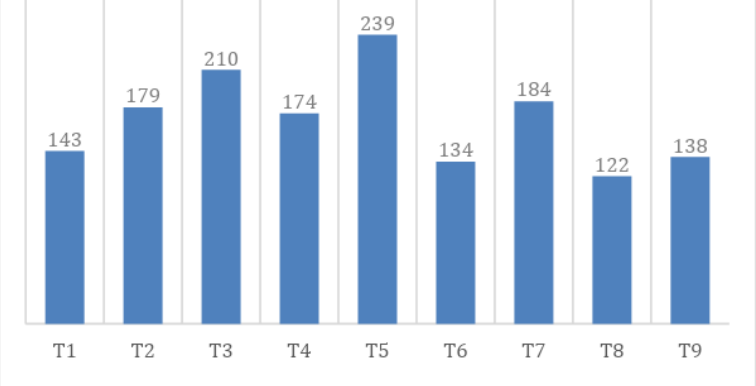
*You can and **should** use your online teaching experience to add new skills to your resume…. (T6, S61)*

Hedges

In line with the test results for the use of boosters, the significant difference in the frequency of hedging devices across tasks led to only one significant pairwise difference, which was found between the eighth task on TPRS with the lowest median frequency of this category and the fifth task on DIY in ELT with one of the highest frequencies observed (Fig. 4). This disparity could once again be attributed to differences in the content and style of the readings. The TPRS papers introduced an approach with which the participants were not familiar, whereas the DIY papers offered practical guidelines that were of great relevance to material design and instructional practices. Therefore, it is likely that the participants found the DIY papers more relatable to their context and discussed potential advantages and challenges for teachers in greater detail. Furthermore, the decline in the frequency of hedges, especially in the final two tasks, contrasts with the findings of Crosthwaite and Jiang (2017), Gürsoy (2023), and Martin-Laguna (2023), in which a gradual increase in the frequency of hedges in undergraduate writing was reported. However, it aligns with Carroll's (2007) analysis of student essays after a critical thinking course, which she interpreted as a decrease in assertiveness resulting from heightened familiarity with critical thinking. Nevertheless, the relatively high frequency of hedges across the tasks, despite noticeable fluctuations, can also be associated with effective persuasion, and stance expression (Bruce,

2016; Lee & Deakin, 2016).

Figure 4.
Median Frequencies of Hedges Across Tasks (per 10,000 words)



*On the other hand, I **feel** that integrating translanguaging into language classrooms would make the point of language learning quite absurd and meaningless. (T4, S32)*

*Therefore, students in primary school **may** find it difficult to adopt a new strategy (T4, S66)*

*My thoughts **might** be affected by me being a student, I hope I won't swallow my words in the future. (T5, S80)*

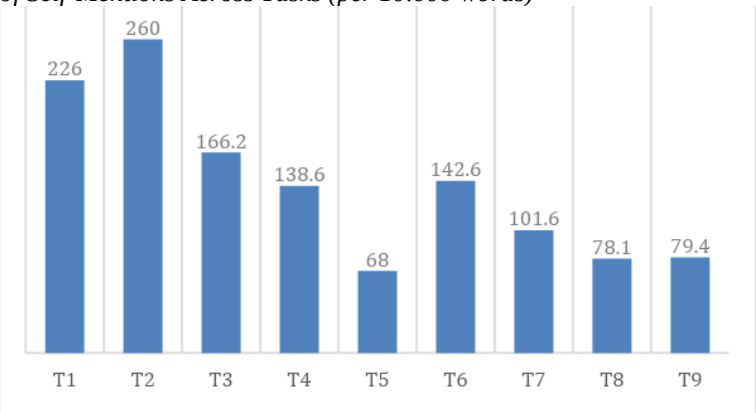
*These **would** be very helpful for acceptance of the other cultures and also diverse students themselves by their teachers and classmates. (T7, S76)*

*Otherwise, the students **could** easily become bored and lose their attention to the class. (T8, S76)*

Self-Mentions

Finally, the last significant post hoc pairwise differences in the use of self-mention devices across tasks, particularly first-person pronouns, point to a remarkable contrast between the first task and the last four tasks (Fig. 5). Although not linear, the use of self-mention devices in the first three tasks continued to decline noticeably. A similar trajectory was observed in Ruan's (2019) analysis of EFL essays by Chinese writers. Having also reported decreased use of personal pronouns after a critical thinking course, Carroll (2007) similarly suggests that this finding may show participants' greater familiarity with evidence-based academic writing conventions towards the end of the semester.

Figure 5.
Median Frequencies of Self-Mentions Across Tasks (per 10.000 words)



***I** think we can say that once the teachers and learners receive necessary trainings on how to correctly use these technologies (T1, S65)*

***I** find it harder to concentrate when reading something from the phone or computer rather than reading from a book. (T1, S11)*

*In this paper **I** will discuss the advantages and disadvantages of teacher-generated materials…. (T5, S75)*

*In **my** opinion, advocacy is not a fixed approach that teachers can apply to all EL. (T7, S61)*

*I made an assumption from that memory of **mine** and read the article carefully. (T3, S80)*

DISCUSSION

This research investigated the use of interactional metadiscourse markers within Hyland's (2005a) metadiscourse model, aiming to provide insights into how critical thinking skills evolve among Turkish pre-service teachers throughout the completion of nine written assignments during a semester. The essays served as the initial stages and tools for exploring and analyzing the subject matter in a critical way. Following Bruce's (2018) argument, this study utilized interactional metadiscourse as a tool to monitor how critical thinking is reflected at the text level. The findings pointed to a complex trajectory alongside noteworthy longitudinal changes in the use of the target markers. While a gradual decrease is evident in the use of the already-infrequent self-mentions and attitude markers towards the end of the semester, as was found in Crosthwaite and Jiang (2017) following explicit MDM instruction, the overall longitudinal development of interactional MDMs was largely not linear, highlighting the complexity of MDM use and the influence of factors such as the writing prompt. The findings offer diverse insights into the understanding of the dynamics of MDM use and its relationship to critical thinking in the context of effective academic writing practices among students, as well as into the instructional design of EAP education.

The results suggest that the students used engagement markers, hedges, and boosters frequently across tasks, while the use of self-mentions and attitude markers decreased towards the end of the semester. In the first task, the most common MDMs were engagement markers, self-mentions, hedges, boosters, and attitude markers, respectively. As of the fourth task, the distribution changed while the most common MDM type remained stable. The order of frequency from the fourth to the ninth task was as engagement markers, hedges, boosters, self-mentions, and attitude markers. The overall use of engagement markers was twice as high as the use of hedges. The prevalent use of engagement markers, which also aligns with previous researches (Beyazyildirim & Ercan, 2023; Ho & Li, 2018; Tasso, 2020; Yüksel & Kavanoz, 2018), suggests a persistent effort by students to engage the reader in their writing, potentially reflecting a strategy learned throughout the course (Tasso, 2020). Additionally, in Crosthwaite and Jiang (2017), in which students received explicit instruction on MDMs in an EAP course, the use of hedging increased, while there was a decrease in the use of boosters and self-mentions. The students tended to use an impersonal and detached writing style, distancing themselves from a more assertive writer identity in academic writing over time. These results suggest that providing explicit instruction on MDMs may not yield radically different results, as somewhat similar findings were obtained in the current research. Therefore, in light of these findings, it can be concluded that relatively extensive exposure to discipline-specific academic texts also has the potential to familiarize students with interactional MDMs to a considerable extent.

Engagement markers, hedges, and boosters were observed to have quite high frequencies regardless of tasks, which might be explained in several ways. As students were new to academic writing and disciplinary conventions, they may have used markers that are relatively more common in colloquial

writing (Qin & Uccelli, 2019). Another reason could be that CRPs allowed students to be more flexible with the use of personal and informal features, and the course instructor did not discourage the use of these informal qualities of writing. As research suggests, academic writing is also slowly becoming less formal (Hyland & Jiang, 2017). Additionally, the finding that the use of hedging devices in particular remained generally high across tasks could be linked to two factors: (1) students' potential attempt at balancing out authorial presence when making strong statements (Yoon, 2021); and (2) the need to adopt a cautious style due to a lack of familiarity with the reading topics (Liu & Stapleton, 2018). In addition, the employment of hedges is also regarded as an indication of high reasoning skills, since it demonstrates that writers take into account the audience's interpretations and perceptions as well (Liu & Stapleton, 2018).

Lower frequencies of self-mentions and attitude markers towards the end of the semester are in line with Carroll's (2007) analysis of papers from a critical thinking course. The students whose papers were analyzed in this study moved from sharing their personal experiences to engaging with ideas in research more closely. This finding is also in line with that of Ruan (2019) in terms of self-mentions. In his study, the less frequent use of self-mentions by Chinese students in EAP essays was associated with the instructions they received about academic writing. In other words, as Hyland (2002) mentioned, L2 students are commonly instructed to write in an objective and impersonal way in academic papers. Likewise, in Lee and Deakin (2016), L2 writers specifically abstained from using self-mentions. Therefore, the students in the study may have adopted a similar standpoint through their high exposure to academic papers throughout the semester.

Consistent with previous studies on Turkish students' essays, the findings of the current study suggest that Turkish students commonly employ hedges in their writing (Algı, 2012; Bayyurt, 2010). Similarly, Çandarlı et al. (2015) mentioned that Turkish students demonstrated a lower level of authorial presence in their English essays compared to their essays written in Turkish. Furthermore, compatible with the current study, Yüksel and Kavanoz (2018) observed that engagement markers were commonly used as interactional MDMs. However, their finding that hedges were underused did not align with the findings of the present research. On the contrary, the students consistently utilized hedges to lower the intensity of their voice as writers. Furthermore, the present investigation could not identify any excessive utilization of self-mentions except for the CRPs at the beginning of the semester.

Additionally, it was observed that the utilization of interactional MDMs exhibited a strong connection with the subject matter of the readings. Students tended to employ these markers more frequently when assessing readings related to instructional methods, subjects they were familiar with, and arguments with which they agreed. For instance, as the students were not familiar with or convinced of the practicality of teaching foreign languages through storytelling (Task 8), limited use of engagement markers, hedges, and boosters was observed in the response papers. These results are compatible with Liu and Stapleton's (2018) and Yoon's (2021) findings that writing prompt selection played a significant role in the usage of interactional MDMs. Therefore, it is plausible to suggest that students also adjust their writing style to the nature of different writing tasks.

CONCLUSION

The purpose of this study was to examine how ELT undergraduate students' writing demonstrates critical thinking through the use of interactional MDMs. In summary, by integrating the fields of ELT, applied linguistics, and corpus linguistics, the study provided insights into the linguistic and textual characteristics of critical writing and the progress of college-level students over a semester. The fluctuations in and distinct trajectories of interactional MDMs in our findings show that, as Ruan (2019) stated, "measuring the developmental trajectory of metadiscourse in university student writing is a complex and multifaceted issue" (p. 484). Further, the findings manifested that topic selection played a seemingly crucial role in the use of interactional MDMs in response papers. Familiarity with and prior

knowledge on the topic, as well as its applicability in the instructional setting, might have led to a more engaged, personal, and strong stance expression. Lastly, it was also clear from the generally fluctuating frequencies of MDMs across tasks that metadiscourse use was rather varied among the study participants, who did not receive any formal instruction on how to effectively use MDMs in their writing.

As for pedagogical implications, several changes can be made to the structure of the course and the writing assignments to include instruction on the effective use of interactional MDMs. As Hyland (2017) states, such instruction has the potential not only to increase students' awareness of the interaction in academic writing but also to improve their overall writing performance. Although the study indicates that extensive exposure to MDMs implicitly allows students to use them to some extent, for more systematic use, explicit instruction is necessary. Therefore, academic writing students can be offered more comprehensive support and facilitation before and after writing CRPs, the kind of writing that necessitates criticality and a vivid expression of personal stance. For instance, as was done in Crosthwaite and Jiang's (2017) study, including instruction on how to express an authorial stance in academic writing, activities, and feedback on the use of interactional MDMs could be integrated into the course plan. The in-class discussions of the readings should also allocate time to further familiarize students with the disciplinary uses of interactional MDMs. Short excerpts from other essential texts in the field could be shared with students, and a reading list including further sources to read on the issues discussed could be provided for students interested in learning more about the topic.

While the study yielded important findings regarding the use of interactional MDMs in non-native undergraduate students' texts, it is not without its limitations, which could be addressed in future studies. Only the writing samples of 22 students who completed all tasks were included in this study, which led to the compilation of a somewhat small corpus. A larger dataset would have allowed findings with greater generalizability. In terms of the instructional setting, the students were not asked to strictly follow a certain writing style or organization. Thus, unlike the corpus projects with detailed compilation guidelines such as the International Corpus of Learner English (ICLE), the CRPs included written output shaped partly by students' personal choices of expression. Another limitation is that this study focused only on the MDMs listed in Hyland (2005a). Going beyond this list by combining it with MDMs included in other sources, manual annotation of our data through corpus-based content analysis could have yielded further findings unique to CRPs. In addition, collecting demographic information from the writers and interviewing them about their views on critical writing could have helped us add a qualitative and contextual dimension to the interpretation of the corpus findings. Lastly, it would have been interesting to also look into the textual borrowing practices of the students, as they frequently made use of the texts they read in their writings. Taking these limitations into account could enable future researchers to strengthen their research designs and add to our understanding of metadiscourse use in critical writing.

REFERENCES

Ädel, A. (2006). *Metadiscourse in L1 and L2 English* (Vol. 24). John Benjamins Publishing.

Aijmer, K., & Rühlemann, C. (2015). *Corpus pragmatics: A handbook.* Cambridge University Press.

Akbaş, E. (2014). Are they discussing in the same way? Interactional metadiscourse in Turkish writers' texts. In *Occupying niches: Interculturality, cross-culturality and aculturality in academic research* (pp. 119-133). Springer.

Algı, S. (2012). *Hedges and boosters in L1 and L2 argumentative paragraphs: Implications for teaching L2 academic writing.* (Publication No. 321188) [MA Thesis, Middle East Technical University], Ankara, Turkey. Council of Higher Education Thesis Center.

Anthony, L. (2017). EncodeAnt (Version 1.2.1) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Anthony, L. (2020). AntConc (Version 3.5.9) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Bayyurt, Y. (2010). Author positioning in academic writing. In S. Zyngier, & V. Viana (Eds.), *Avaliaçoes e perspectivas: mapeando os estudos empiricos na area de Humanas (Appraisals and perspectives: Mapping empirical studies in the humanities)* (pp. 163-184). Publit.

Beyazyildirim, D., & Ercan, G. S. (2023). Interactional Metadiscourse Markers in Argumentative Texts of English Preparatory Class Students. *Journal of Academic Social Science Studies*, *16*(98).

Breeze, R. (2012). *Rethinking academic writing pedagogy for the European university.* Rodopi.

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

Bruce, I. (2016). Constructing critical stance in university essays in English literature and sociology. *English for Specific Purposes*, *42*, 13-25.

Bruce, I. (2018). The textual expression of critical thinking in PhD discussions in Applied Linguistics. *ESP Today*, *6*(1), 2-24. https://doi.org/10.18485/esptoday.2018.6.1.1

Carroll, D. W. (2007). Patterns of student writing in a critical thinking course: A quantitative analysis. *Assessing Writing*, *12*(3), 213-227. https://doi.org/10.1016/j.asw.2008.02.001

Condon, W., & Kelly-Riley, D. (2004). Assessing and teaching what we value: The relationship between college-level writing and critical thinking abilities. *Assessing Writing*, 9, 56–75. https://doi.org/10.1016/j.asw.2004.01.003

Conover, W. J. (1999). *Practical nonparametric statistics* (Vol. 350). John Wiley & Sons.

Crosthwaite, P., & Jiang, K. (2017). Does EAP affect written L2 academic stance? A longitudinal learner corpus study. *System*, *69*, 92-107. https://doi.org/10.1016/j.system.2017.06.010

Çandarlı, D., Bayyurt, Y., & Martı, L. (2015). Authorial presence in L1 and L2 novice academic writing: Cross-linguistic and cross-cultural perspectives. *Journal of English for Academic Purposes*, *20*, 192-202. https://doi.org/10.1016/j.jeap.2015.10.001

Dahl, T. (2004). Textual metadiscourse in research articles: a marker of national culture or of academic discipline?. *Journal of Pragmatics*, *36*(10), 1807-1825. https://doi.org/10.1016/j.pragma.2004.05.004

Ennis, R. (1987). A taxonomy of critical thinking abilities and dispositions. In J. Baron and R. Sternberg (Eds.), *Teaching thinking skills: Theory and practice*. (pp. 9-26). W.H. Freeman.

Ennis, R. (1992). The degree to which critical thinking is subject specific: Clarification and needed research. In S. Norris (Ed.), *The generalizability of critical thinking: Multiple perspectives on an educational ideal* (pp. 21-37). Teachers College Press.

Facione, P. A. (2011). *Critical thinking: What it is and why it counts*. California Academic Press.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2023). *How to design and evaluate research in education* (11th ed.). McGraw Hill.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, *32*, 674–701. https://doi.org/10.1080/01621459.1937.10503522

Gillaerts, P., & Van de Velde, F. (2010). Interactional metadiscourse in research article abstracts. *Journal of English for Academic Purposes*, *9*(2), 128-139. https://doi.org/10.1016/j.jeap.2010.02.004

Gürsoy, Z. (2023). *The investigation of metadiscourse markers in native and non-native academic corpora: Authorial stance and the use of hedging.* (Publication No. 839408) [MA Thesis, Karadeniz Technical University], Trabzon, Turkey. Council of Higher Education Thesis Center.

Ho, V., & Li, C. (2018). The use of metadiscourse and persuasion: An analysis of first year university students' timed argumentative essays. *Journal of English for Academic Purposes*, *33*, 53-68.

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of pragmatics*, *34*(8), 1091-1112.

Hyland, K. (2004). Patterns of engagement: Dialogic features and L2 undergraduate writing. In L. Ravelli & R. A. Ellis (Eds.), *Analysing academic writing: Contextualized frameworks* (pp. 5-23). Continuum.

Hyland, K. (2005a). *Metadiscourse: Exploring interaction in writing*. A&C Black.

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

Hyland, K. (2013). Faculty feedback: Perceptions and practices in L2 disciplinary writing. *Journal of Second Language Writing*, *22*(3), 240-253. https://doi.org/10.1016/j.jslw.2013.03.003

Hyland, K. (2017). Metadiscourse: What is it and where is it going?. *Journal of Pragmatics*, *113*, 16-29. https://doi.org/10.1016/j.pragma.2017.03.007

Hyland, K., & Jiang, F. K. (2018). "In this paper we suggest": Changing patterns of disciplinary metadiscourse. *English for Specific Purposes*, *51*, 18-30. https://doi.org/10.1016/j.esp.2018.02.001

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, *25*(2), 156-177. https://doi.org/10.1093/applin/25.2.156

Koester, A. (2010). Building small specialised corpora. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 66- 79). Routledge.

Lancaster, Z. (2016). Expressing stance in undergraduate writing: Discipline-specific and general qualities. *Journal of English for Academic Purposes*, *23*, 16-30.

Lee, J. J., & Deakin, L. (2016). Interactions in L1 and L2 undergraduate student writing: Interactional metadiscourse in successful and less-successful argumentative essays. *Journal of Second Language Writing*, *33*, 21-34.

Liu, F., & Stapleton, P. (2018). Connecting writing assessment with critical thinking: An exploratory study of alternative rhetorical functions and objects of enquiry in writing prompts. *Assessing Writing*, *38*, 10-20. https://doi.org/10.1016/j.asw.2018.09.001

Martín-Laguna, S. (2023). Metadiscourse learning trajectories in multilingual learners: A focus on attitude markers and hedges. In *New trends on metadiscourse: An analysis of online and textual genres* (pp. 105-128). Springer International Publishing.

McKinley, J. (2013). Displaying critical thinking in EFL academic writing: A discussion of Japanese to English contrastive rhetoric. *RELC Journal*, *44*(2), 195-208. https://doi.org/10.1177/0033688213488386

Moore, T. (2013). Critical thinking: Seven definitions in search of a concept. *Studies in Higher Education*, *38*(4), 506-522. https://doi.org/10.1080/03075079.2011.586995

Nosich, G. (2021). *Critical writing: a guide to writing a paper using the concepts and processes of critical thinking*. Rowman & Littlefield.

Pohlert, T. (2018). *PMCMR: Calculate Pairwise Multiple Comparisons of Mean Rank Sums*. [R package]. Retrieved from https://cran.r-project.org/package=PMCMR.

Qin, W., & Uccelli, P. (2019). Metadiscourse: Variation across communicative contexts. *Journal of Pragmatics*, *139*, 22-39. https://doi.org/10.1016/j.pragma.2018.10.004

Ross, D. (2018). Small corpora and low-frequency phenomena: try and beyond contemporary, standard English. *Corpus*, *18*, 1-40. https://doi.org/10.4000/corpus.3574

Ruan, Z. (2019). Metadiscourse use in L2 student essay writing: A longitudinal cross-contextual comparison. *Chinese Journal of Applied Linguistics*, *42*(4), 466-487. https://doi.org/10.1515/CJAL-2019-0028

Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students: Essential tasks and skills* (3rd ed.). University of Michigan Press. https://doi.org/10.3998/mpub.2173936

Tasso, C. (2020). Interactional metadiscourse features in the written production made by Spanish EFL learners. In M. L. Carrió-Pastor (Ed.), *Corpus analysis in different genres: Academic discourse and learner corpora* (pp. 206-218). Routledge.

The Jamovi Project (2021). *Jamovi*. (Version 2.2) [Computer Software]. Retrieved from https://www.jamovi.org.

Thompson, G., & Thetela, P. (1995). The sound of one hand clapping: The management of interaction in written discourse. *Text & Talk*, *15*(1), 103-128. https://doi.org/10.1515/text.1.1995.15.1.103

Trillo, J. R. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics*, *34*(6), 769-784. https://doi.org/10.1016/S0378-2166(02)00022-X

Woodward-Kron, R. (2002). Critical analysis versus description? Examining the relationship in successful student writing. *Journal of English for Academic Purposes*, *1*(2), 121-143. https://doi.org/10.1016/S1475-1585(02)00013-9

Yoon, H. J. (2021). Interactions in EFL argumentative writing: effects of topic, L1 background, and L2 proficiency on interactional metadiscourse. *Reading and Writing*, *34*(3), 705-725. https://doi.org/10.1007/s11145-020-10085-7

Yüksel, H. G., & Kavanoz, S. (2018). Dimension of Experience: Metadiscourse in the Texts of Novice Non-Native, Novice Native and Expert Native Speaker. *Advances in language and literary studies*, *9*(3), 104-112.

APPENDICES

Appendix 1: Details of Weekly Discussions

Theme 1 - New Digital Platforms

- Alhumaid, K. (2019). Four ways technology has negatively changed education. *Journal of Educational and Social Research*, *9*(4), 10-10.

- Hockly, N., & Dudeney, G. (2018). Current and future digital trends in ELT. *Relc Journal*, *49*(2), 164-178.

Theme 2 - Mindfulness and Social-Emotional Teaching

- Elias M. J. (2018) *Pros and Cons of Mindfulness in SEL*. Retrieved from *https://www.edutopia.org/article/pros-and-cons-mindfulness-sel*

- Johnson, K. E., & Golombek, P. R. (2016). *Mindful L2 teacher education: A sociocultural perspective on cultivating teachers' professional development*. Routledge.

- Pentón Herrera, L. J. (2020). Social-emotional learning in TESOL: What, why, and how. *Journal of English Learner Education*, *10*(1), 1.

Theme 3 - What Do Successful Teachers and Learners Do?

- Pasini, V. (2017, June 5), Cambridge Papers #1: What do successful language teachers and learners do?. Cambridge ELT Blog. Retrieved from https://www.cambridge.org/elt/blog/2017/06/05/cambridge-papers-1-successful-language-teachers-learners/

- Uygun, S. (2013). How to become an effective English language teacher. *Journal of Educational and Social Research*, *3*(7), 306-311.

Theme 4 - Translanguaging

- Wang R. (2018): New perspectives on translanguaging and education. *International Journal of Bilingual Education and Bilingualism,* 1-3. DOI: 10.1080/13670050.2018.1454043

- What is Translanguaging? (2016, July 26) NALDIC's EAL Blog. Retrieved from *https://ealjournal.org/2016/07/26/what-is-translanguaging/*

Theme 5 - DIY in ELT

- Howard, J., & Major, J. (2004). Guidelines for designing effective English language teaching materials. *The TESOLANZ Journal*, *12*(10), 50-58.

- Wyatt, M. (2011). Becoming a do-it-yourself designer of English language teaching materials. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, *12* (1), 1-38.

Theme 6 - Online ESL/EFL Classes

- Mutonono E. (blog) Retrieved from *https://www.teachingeslonline.com/advantages-challenges-online-elena-mutonono/* in September 2020.

- Pachina E. (2019). Advantages and Disadvantages of Online Education in The ESL Classroom. Retrieved from *https://www.teflcourse.net/blog/advantages-and-disadvantages-of-online-education-in-the-esl-classroom-ittt-tefl-blog/*

Theme 7 - Advocacy in ELT

- Linville, H. A., & Whiting, J. (Eds.). (2019). *Advocacy in English language teaching and learning*. Routledge.

- Students chose and read one more article on the topic.

Theme 8 - Teaching Proficiency Through Reading and Storytelling (TPRS)

- Octaviani, A., Hesmatantya, V., & DediWijaya, S. (2018). Using Teaching Proficiency Through Reading and Storytelling (TPRS) in Teaching English for Young Learners, *Tell Journal, 6* (2), 78-87

- Pardede, P. (2011). Using short stories to teach language skills. *JET (Journal of English Teaching), 1*(1), 14-27.

Theme 9 - Pros and Cons of Raising Bilingual Children

- Pros and Cons of Raising Bilingual Children (N.D.) Raising Bilingual Children Blog. Retrieved from *https://www.raising-bilingual-children.com/basics/info/pros-cons/* in September 2020.

- Students were asked to choose and read two articles regarding the pros and cons of raising bilingual children, then write a critical review on the subject. The blog post above was given as a starting point.

Appendix 2: 10 Most Frequent Interactional Metadiscourse Markers

Hedges

	Task 1			Task 2			Task 3		
No	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*would*	53	24.32	*may*	29	15.23	*may*	22	11.00
2	*may*	42	19.27	*feel*	24	12.60	*rather*	20	10.00
3	*might*	22	10.10	*would*	20	10.50	*would*	20	10.00
4	*feel*	19	8.72	*claim*	13	6.83	*feel*	19	9.50
5	*in my opinion*	17	7.80	*could*	11	5.78	*could*	17	8.50
6	*could*	16	7.34	*might*	11	5.78	*might*	16	8.00
7	*claim*	14	6.42	*argue*	9	4.73	*claim*	11	5.50
8	*possible*	14	6.42	*sometimes*	9	4.73	*possible*	11	5.50
9	*quite*	8	3.67	*suggest*	8	4.20	*indicate*	10	5.00
10	*assume*	7	3.21	*often*	7	3.68	*suggest*	10	5.00
11	*mostly*	7	3.21	*possible*	7	3.68	*in my opinion*	9	4.50
12	*seem*	7	3.21	*in my opinion*	6	3.15	*argue*	6	3.00

No	Task 4			Task 5			Task 6		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*may*	60	39.87	*may*	73	37.01	*may*	39	24.03
2	*would*	33	21.93	*would*	50	25.35	*might*	20	12.32
3	*could*	20	13.29	*could*	20	10.14	*could*	19	11.71
4	*feel*	19	12.63	*might*	20	10.14	*possible*	17	10.47
5	*might*	19	12.63	*around*	12	6.08	*would*	14	8.63
6	*in my opinion*	9	5.98	*feel*	11	5.58	*in my opinion*	11	6.78
7	*possible*	8	5.32	*possible*	10	5.07	*feel*	8	4.93
8	*mostly*	6	3.99	*in my opinion*	8	4.06	*indicate*	5	3.08
9	*often*	6	3.99	*usually*	8	4.06	*quite*	5	3.08
10	*quite*	6	3.99	*mostly*	7	3.55	*claim*	4	2.46
11	*rather*	6	3.99	*often*	7	3.55	*likely*	4	2.46
12	*argue*	4	2.66	*seem*	7	3.55	*seem*	4	2.46

No	Task 7			Task 8			Task 9		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*feel*	39	19.61	*may*	25	14.81	*may*	44	20.47
2	*may*	30	15.08	*could*	18	10.66	*would*	25	11.63
3	*could*	30	15.08	*would*	14	8.29	*might*	18	8.37
4	*would*	26	13.07	*in my opinion*	9	5.33	*could*	17	7.91
5	*might*	20	10.06	*possible*	9	5.33	*feel*	15	6.98
6	*suggest*	12	6.03	*feel*	6	3.55	*claim*	11	5.12
7	*in my opinion*	11	5.53	*indicate*	6	3.55	*indicate*	9	4.19
8	*possible*	11	5.53	*mostly*	6	3.55	*argue*	8	3.72
9	*likely*	8	4.02	*usually*	6	3.55	*in my opinion*	8	3.72
10	*sometimes*	8	4.02	*almost*	4	2.37	*tend to*	8	3.72
11	*often*	6	3.02	*quite*	4	2.37	*often*	7	3.26
12	*argue*	5	2.51	*argue*	3	1.78	*possible*	7	3.26

Self-Mentions

No	Task 1			Task 2			Task 3		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*I*	205	94.08	*I*	121	63.54	*I*	137	68.48
2	*my*	68	31.21	*my*	31	16.28	*my*	41	20.49
3	*me*	31	14.23	*me*	8	4.20	*me*	18	9.00
4							*mine*	1	0.50

No	Task 4			Task 5			Task 6		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*I*	102	67.78	*I*	118	59.82	*I*	103	63.46
2	*my*	34	22.59	*my*	32	16.22	*my*	24	14.79
3	*me*	9	5.98	*me*	8	4.06	*me*	4	2.46

No	Task 7			Task 8			Task 9		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*I*	100	50.28	*I*	93	55.08	*I*	104	48.38
2	*my*	26	13.07	*my*	25	14.81	*my*	28	13.02
3	*me*	10	5.03	*me*	8	4.74	*me*	6	2.79

Boosters

No	Task 1			Task 2			Task 3		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*think*	53	24.32	*think*	43	22.58	*think*	36	18.00
2	*believe*	24	11.01	*know*	23	12.08	*find*	34	17.00
3	*find*	19	8.72	*must*	21	11.03	*know*	23	11.50
4	*know*	18	8.26	*find*	20	10.50	*must*	23	11.50
5	*really*	17	7.80	*believe*	19	9.98	*show*	20	10.00
6	*always*	13	5.97	*always*	15	7.88	*believe*	13	6.50
7	*show*	10	4.59	*true*	10	5.25	*clearly*	13	6.50
8	*clearly*	8	3.67	*certain*	6	3.15	*always*	12	6.00
9	*must*	7	3.21	*clearly*	5	2.63	*clear*	12	6.00
10	*actually*	4	1.84	*demonstrate*	5	2.63	*really*	10	5.00

No	Task 4			Task 5			Task 6		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*think*	33	21.93	*think*	34	17.24	*think*	38	23.41
2	*know*	22	14.62	*must*	27	13.69	*find*	14	8.63
3	*find*	14	9.30	*find*	22	11.15	*always*	12	7.39
4	*believe*	13	8.64	*clear*	21	10.65	*know*	12	7.39
5	*show*	12	7.97	*believe*	16	8.11	*believe*	11	6.78
6	*must*	9	5.98	*know*	15	7.60	*must*	11	6.78
7	*really*	9	5.98	*of course*	9	4.56	*of course*	6	3.70
8	*clear*	8	5.32	*clearly*	8	4.06	*really*	6	3.70
9	*always*	7	4.65	*certain*	7	3.55	*realize*	4	2.46
10	*certain*	7	4.65	*always*	6	3.04	*show*	4	2.46

No	Task 7			Task 8			Task 9		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*know*	29	14.58	*think*	21	12.44	*think*	40	18.61
2	*think*	28	14.08	*find*	15	8.88	*show*	30	13.95
3	*must*	25	12.57	*show*	15	8.88	*believe*	23	10.70
4	*find*	14	7.04	*believe*	12	7.11	*know*	20	9.30
5	*believe*	13	6.54	*must*	9	5.33	*find*	18	8.37
6	*always*	12	6.03	*certain*	8	4.74	*must*	11	5.12
7	*show*	8	4.02	*really*	7	4.15	*actually*	8	3.72
8	*really*	6	3.02	*establish*	6	3.55	*always*	7	3.26
9	*sure*	6	3.02	*true*	6	3.55	*clear*	7	3.26
10	*actually*	5	2.51	*actually*	5	2.96	*certain*	6	2.79

Engagement Markers

No	Task 1			Task 2			Task 3		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*we*	102	46.81	*we*	158	82.97	*should*	197	98.48
2	*use*	91	41.76	*should*	130	68.27	*use*	54	26.99
3	*our*	46	21.11	*our*	92	48.31	*we*	45	22.49
4	*should*	36	16.52	*use*	34	17.86	*?*	31	15.50
5	*do not*	35	16.06	*?*	23	12.08	*you*	25	12.50
6	*us*	22	10.10	*need to*	23	12.08	*find*	23	11.50
7	*does not*	20	9.18	*must*	21	11.03	*input*	23	11.50
8	*take*	20	9.18	*take*	20	10.50	*must*	23	11.50
9	*find*	17	7.80	*us*	20	10.50	*does not*	16	8.00
10	*see*	16	7.34	*you*	20	10.50	*do not*	15	7.50

No	Task 4			Task 5			Task 6		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*use*	88	58.48	*should*	184	93.28	*we*	62	38.20
2	*should*	53	35.22	*use*	66	33.46	*do not*	55	33.89
3	*we*	44	29.24	*consider*	27	13.69	*you*	40	24.65
4	*you*	26	17.28	*develop*	27	13.69	*should*	37	22.80
5	*review*	23	15.28	*must*	27	13.69	*take*	34	20.95
6	*?*	20	13.29	*take*	26	13.18	*have to*	27	16.64
7	*do not*	19	12.63	*choose*	25	12.67	*our*	27	16.64
8	*does not*	17	11.30	*do not*	23	11.66	*your*	24	14.79
9	*develop*	13	8.64	*prepare*	21	10.65	*does not*	19	11.71
10	*your*	13	8.64	*we*	20	10.14	*use*	17	10.47

No	Task 7			Task 8			Task 9		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*should*	141	70.90	*use*	69	40.87	*should*	64	29.77
2	*we*	63	31.68	*should*	58	34.35	*use*	36	16.75
3	*do not*	28	14.08	*we*	23	13.62	*you*	35	16.28
4	*must*	25	12.57	*do not*	11	6.52	*do not*	34	15.82
5	*take*	21	10.56	*need to*	11	6.52	*we*	34	15.82
6	*our*	20	10.06	*our*	11	6.52	*your*	28	13.02
7	*need to*	19	9.55	*us*	11	6.52	*?*	27	12.56
8	*you*	18	9.05	*develop*	10	5.92	*does not*	20	9.30
9	*?*	16	8.05	*take*	10	5.92	*develop*	18	8.37
10	*see*	13	6.54	*?*	9	5.33	*need to*	15	6.98

Attitude Markers

No	Task 1			Task 2			Task 3		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*even*	44	20.19	*important*	31	16.28	*agree*	45	22.49
2	*agree*	32	14.68	*agree*	23	12.08	*important*	37	18.50
3	*important*	18	8.26	*even*	20	10.50	*even*	19	9.50
4	*interesting*	10	4.59	*essential*	13	6.83	*disagree*	17	8.50
5	*appropriate*	6	2.75	*appropriate*	4	2.10	*essential*	8	4.00
6	*disagree*	5	2.29	*interesting*	3	1.58	*interesting*	5	2.50
7	*essential*	3	1.38	*unfortunately*	3	1.58	*understandable*	5	2.50
8	*dramatic*	2	0.92	*prefer*	2	1.05	*desirable*	3	1.50
9	*expected*	2	0.92	*appropriately*	1	0.53	*appropriate*	2	1.00
10	*prefer*	2	0.92	*disagree*	1	0.53	*expected*	2	1.00

No	Task 4			Task 5			Task 6		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*even*	21	13.96	*important*	38	19.26	*even*	32	19.72
2	*agree*	14	9.30	*appropriate*	19	9.63	*agree*	25	15.40
3	*important*	8	5.32	*agree*	18	9.13	*important*	10	6.16
4	*appropriate*	2	1.33	*even*	17	8.62	*appropriate*	6	3.70
5	*essential*	2	1.33	*essential*	10	5.07	*prefer*	4	2.46
6	*correctly*	1	0.66	*interesting*	9	4.56	*!*	3	1.85
7	*desirable*	1	0.66	*prefer*	6	3.04	*essential*	3	1.85
8	*hopeful*	1	0.66	*understandable*	5	2.53	*unfortunately*	3	1.85
9	*prefer*	1	0.66	*inappropriate*	3	1.52	*importantly*	2	1.23
10	*understandable*	1	0.66	*unfortunately*	2	1.01	*unexpectedly*	2	1.23

No	Task 7			Task 8			Task 9		
	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)	Item	f(raw)	f(normed)
1	*important*	47	23.63	*important*	16	9.48	*even*	26	12.09
2	*even*	19	9.55	*even*	14	8.29	*important*	19	8.84
3	*agree*	11	5.53	*agree*	12	7.11	*agree*	18	8.37
4	*appropriate*	6	3.02	*appropriate*	12	7.11	*prefer*	9	4.19
5	*essential*	4	2.01	*interesting*	5	2.96	*!*	4	1.86
6	*appropriately*	3	1.51	*prefer*	5	2.96	*expected*	4	1.86
7	*understandable*	2	1.01	*essential*	4	2.37	*essential*	3	1.40
8	*expected*	1	0.50	*correctly*	2	1.18	*correctly*	1	0.47
9	*striking*	1	0.50	*importantly*	2	1.18	*disagree*	1	0.47
10	*unfortunately*	1	0.50	*unexpected*	2	1.18	*disappointing*	1	0.47

TÜRKÇE GENİŞLETİLMİŞ ÖZET

Üniversite düzeyinde alana özgü bilgiyi dilbilgisel ve yapısal olarak uygun bir şekilde iletmek ana dili İngilizce olmayan yazarlar için bilişsel zorluklar barındırmaktadır (Breeze, 2012; Hyland, 2013). Yazarlar belirli iletişim sinyallerini kullanmaya başladıklarında, metindeki ileti okuyucu tarafından daha verimli bir şekilde değerlendirilebilir. Dolayısıyla, "bir metni ya da yazarın tutumunu, içeriğe ya da okuyucuya göre düzenleyen metin yönlendiricileri"(s. 14) anlamına gelen üstsöylem, akademik amaçlar için yazma eğitiminin önemli bileşenlerinden biri olarak kabul edilmektedir (Hyland, 2005a). Üstsöylem belirteçleri, metinlerde yazarların önerilen konuya karşı tutumunu yansıtır ve okuyucuların algısını etkiler (Hyland, 2005a). Yazarların kimlikleri, inançları ve içerikle ilgili tutumları da bu belirteçlerin kullanımı aracılığıyla gözlemlenebilmektedir (Hyland, 2005a; Trillo, 2002).

Üniversite düzeyinde yabancı bir dilde etkili yazmanın önemli bir diğer değişkeni de, yazılı metinlerde eleştirel düşünme becerilerinin yansıtılmasıdır. Yapılan çalışmalar eleştirel düşünme ile yazma becerileri arasında karşılıklı bir ilişki olduğunu göstermiş, üniversite düzeyinde yazı yazmanın öğrencilerin alan bilgilerini arttırdığını ve eleştirel düşünme becerilerini geliştirdiğini kanıtlamıştır (Carroll, 2007).

Eleştirel düşünme becerileri, güçlü argümanların ve etkileşimsel üstsöylem belirteçlerinin (EÜB) etkili kullanımıyla yazıya yansıtılabilmektedir. Bruce (2018) tarafından öne sürüldüğü gibi, özellikle etkileşimsel üstsöylem, eleştirel düşünmenin metin düzeyindeki özelliklerini gözlemlemek için kullanılabilecek analitik bir araçtır. Bu tür araçlar, okuyucuyu ikna etmek, tutumları ve fikirleri etkili bir şekilde ifade etmek için kullanılabilir. Örneğin, EÜB'ler aracılığıyla yazarlar görüşlerini iletebilir (düşünüyorum, hissediyorum) ya da hafifletme belirteçleri kullanarak iddialarının gücünü değiştirebilirler (gerekmektedir, olabilir) (Liu & Stapleton, 2018).

Lisans ve lisansüstü düzeyinde eleştirel yazma üzerine yapılan gelişimsel ve karşılaştırmalı araştırmalar, öğrencilerin yazma sürecine kontrollü belirteçlerin, içgörü ve kendine atıfta bulunma belirteçlerinin sıklıkla kullanıldığı metinlerle başladığını göstermiştir. Gelişimsel süreç içinde ise hafifletme ve güçlendirme belirteçlerinin kullanımı artmaktadır. Bu bulgular, nedensel düşünmenin birer yansıması olarak nitelendirilmektedir. Ayrıca, yazma konuları ve söylem toplulukları, öğrencilerin EÜB'leri kullanımını etkileyebilmektedir. Alanyazındaki bulgular ışığında yürütülen bu çalışma, bir Türk devlet üniversitesinde İngilizce öğretmeni adayları tarafından bir dönem boyunca yazılan eleştirel yanıt makalelerinde kullanılan ve Hyland (2005a) tarafından sınıflandırılan EÜB'ler aracılığıyla eleştirel düşüncenin dile yansımasını araştırmayı amaçlamıştır. Çalışmanın amaçları doğrultusunda, bu araştırma sorularına cevap aranmıştır: (1) İngilizce öğretmen adaylarının eleştirel değerlendirme raporlarında, Hyland (2005a) tarafından belirtilen EÜB'lerden hangileri sıklıkla kullanılmıştır? (2) Eleştirel değerlendirme raporu ödevlerinde EÜB'lerin kullanımı nasıl farklılıklar göstermektedir?

Araştırmada 22 öğrenci tarafından yazılan, 198 yanıt belgesini içeren Eleştirel Yanıt Derlemi (bundan böyle EYD olarak anılacak) incelenmiştir. Katılımcı grubunu İngilizce Öğretmenliği bölümü ikinci sınıf öğrencilerine sunulan "Eleştirel Okuma ve Yazma" dersine kayıtlı öğrenciler oluşturmaktadır. Dönem boyunca öğrenciler, diller arası geçişlilik, İngilizce derslerinde dijital platformların kullanımı, savunuculuk, farkındalık ve sosyo-duygusal öğrenme gibi konular üzerine her hafta sınıfta irdelenen ve tartışılan makalelere yönelik toplamda dokuz eleştirel değerlendirme raporu yazmıştır. Hyland (2005a) tarafından listelenen üstsöylem belirteçlerinin sıklıları, derlem üzerinde AntConc 3.5.9 (Anthony, 2020) kullanılarak saptanmıştır. Çalışma örnekleminin küçük olması (N=22) ve dağılımın normal olmaması nedeniyle Friedman'ın (1937) parametrik olmayan tekrarlı ölçümlerde tek yönlü varyans analizi (ANOVA) ve ardından Durbin-Conover (Conover, 1999) PostHoc testleri, Jamovi 2.2 (2021) istatistik analiz programı kullanılarak yapılmıştır.

Genel olarak, bulgular üniversite düzeyindeki öğrencilerin eleştirel yazımlarındaki dilbilgisel özelliklerine ve yazma becerilerindeki gelişimsel süreçlerine ışık tutmaktadır (Woodward-Kron, 2002).

Örneğin, metinlerde kendine atıfta bulunma ve tutum belirteçlerinin kullanımında azalma gibi anlamlı, boylamsal değişiklikler gözlemlenmiştir. Buna ek olarak, sonuçlar konu seçiminin EÜB'lerin kullanımında önemli bir rol oynadığını göstermektedir.

Çalışma bulguları ışığında, EÜB'lerin etkili kullanımı üzerine çeşitli pedagojik önerilerde bulunulabilir. Örneğin, ders içerikleri ve yazma ödevleri EÜB kullanımını öğretecek ve teşvik edecek şekilde hazırlanabilir. Hyland (2017)'ın vurguladığı gibi, öğrencilerin akademik yazıda etkileşim becerilerini geliştirmenin yanı sıra genel yazma performansları da bu şekilde yükseltilebilir.