SAL 34/2019

Metadiscourse in Academic English Texts: A Corpus-based Probe into British Academic Written English Corpus

Received 10/2018
Accepted 03/2019

Metadiscourse in Academic English Texts: A Corpus-based Probe into British Academic Written English Corpus

Metadiskursas akademiniuose anglų kalbos tekstuose: britų akademinės rašytinės anglų kalbos tekstyno tyrimas

LINGUISTICS / KALBOTYRA

Mehrdad Vasheghani Farahani
Leipzig University, Germany

http://dx.doi.org/10.5755/j01.sal.0.34.21816

Abstract

This paper reports on a research performed in the field of corpus linguistics on metadiscourse features in the British Academic Written English Corpus. For this purpose, the British Academic Written English Corpus, which is freely available and contains 6,968,089 words, was selected as the data resource of the study. The taxonomy of metadiscourse features compiled by Hyland was used as the theoretical framework and the R program was used as the statistical software to run statistical analysis. As the data show, the interactive metadiscourse features were more prevalent than the interactional metadiscourse features in the corpus. In the interactive category, transitions and endophoric markers were used more than other ones; whereas, in the interactional category, hedges and boosters were the predominant metadiscourse features. The prevalence nature of interactive metadiscourse features can add support to the idea that writers were more interested in organising discourse rather than conducting interaction to the audience. The findings of this research can have useful implications for researchers in such fields as contrastive analysis, text linguistics and corpus-based studies.

KEYWORDS: English academic texts, metadiscourse features, interactive and interactional, British Academic Written English Corpus, corpus research.

Introduction

Written language plays a salient role in various aspects of our lives including academic, professional and social ones (Vasheghani Farahani & Sabetifard, 2017). In other words, writing is crucially a significant phenomenon in that it is the basis upon which communication, history, and every other interaction are recorded and/or taken place. Indeed, "writing encompasses creative inspiration, problem-solving, reflection and revision that results in a completed manuscript" (Defazio, Jones, Tennant & Hook, 2010, p. 34). Writing is important in that academic success in every discipline depends, to a large extent, on writing skills (Cho & Schunn, 2007).

Writing can take various genres based on the setting in which it is used and the purposes it seeks. From among various genres, one is academic writing, which shares the idea of many

ktu 1922

Research Journal
Studies about Languages
No. 34/2019
ISSN 1648-2824 (print)
ISSN 2029-7203 (online)
pp. 56-73
DOI 10.5755/j01.sal.34.0.21816

speakers of English that it is peculiar from other genre and registers (Biber & Gray, 2016). In this regard, every genre as a specific text type enjoys specific features, which distinguishes it from other types. For example, a text in the field of news writing is distinct from a piece of writing in the field of literature. According to Swales (1990), "exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience" (p. 58).

The purpose and the receivers of the writing for whom the writing as a specific genre is written are among the factors which can have impacts on the production of writing product (Tadayyon & Vasheghani Farahani, 2017). Therefore, academic writing is different from other types of writing in that its focus is in on "those communication skills in English which are required for study purposes informal education systems" (Jordan, 1997, p. 1). Furthermore, complex grammatical structures and exploiting the explicit language are the two main characterisations of the academic writing in a sense that "academic writing is more complex, structurally elaborated, and explicit in meaning than most other spoken and written registers" (Biber & Gray, 2016, p. 14). In addition, academic writing is determined and detected by longer wording, more precision, and more difficulty to understand (Bailey, 2015).

Academic writing is used for miscellaneous purposes such as reporting the results of a research, answering a research question, discussing a scientific topic and synthesizing researchers done by others (Bailey, 2003). Indeed, "academic writers generate texts as much to represent some external reality as to display their attitudinal positions in relation to the external reality and the recipients thereof" (Zarei & Mansoori, 2007, p. 25). It is a skill which has to be mastered by students and researchers in order to be able to produce research papers, academic books and dissertations. Bowker (2007) mentions some of the most salient features of academic writing. For her, one difference is the application of punctuation and grammar that follows very strict rules. Apart from punctuation and grammar, academic writing focuses mostly on abstract ideas, which cannot be explained in the physical form. In addition, academic writing requires the use of citation and reference to other works. Evidently, non-academic writing does not stick to these rules, making the process of writing easier and less complex.

One of the most important elements for yielding an academic writing production is metadiscourse features. The term *metadiscourse* was first introduced by Harris in 1959 to refer to a new way of understanding the complexities of language in real context and to represent the way writers or speakers guide the course of the message (Hyland, 2005). Also known as non-topical elements (Lautamati, 1979) or discourse about discourse (Hyland, 2002; 2018), metadiscourse refers to an open class of lexical items which can play pragmatically important roles in establishing the interaction between the writer and the reader (Gholami, Tajalli & Shokrpour, 2014).

Metadiscourse takes our attention to "the ways writers project themselves into their discourse to signal their attitude towards both the content and the audience of the text" (Hyland & Tse, 2004). As a result, a central role of metadiscourse is in the genre analysis. In other words, metadiscourse features are highly text/genre dependent (Hyland, 2000). This context dependency becomes apparent when metadiscourse features are used in various genres including academic writing. As a matter of fact, this context dependency of metadiscourse features helps writers "respond to and construct the contexts in which language is used" (Hyland, 2019, p. 104).

Corpus as "a collection of spoken or written texts to be used for linguistic analysis and based on a specific set of design criteria influenced by its purpose and scope" (Weisser, 2016, p. 13) has found its way to language studies in line with advances in computer science. One area of study for exploiting corpus software is metadiscourse features. In fact, by exploiting corpus

software, it is possible to analyse a large number of texts that are thought to contain instances of metadiscourse features. In this regard, analysing metadiscourse features by using large and representative corpora is a good way of delving into the changes of languages in various genres (Boggel, 2008).

Most studies in this field of research suffer from two major shortcomings. One is that these studies are done based on a small scale of data that may question the issue of external validity and generalisability. The other shortcoming is that most of these studies lack the application of corpora; therefore, their methodology can be questioned.

Literature Review

Metadiscourse Definition and Significance

A number of various scholars have defined metadiscourse in different modes, as it is a relatively new area of research (Heng & Tan, 2010). Hyland and Tse (2004) put forward the notion that the definition of metadiscourse as discourse about discourse is wrong, stating that metadiscourse is a kind of linguistic resource by which the author can organise his discourse and his stance towards the receiver of the message. In another definition, metadiscourse features are defined as "a new and interesting field of inquiry which is believed to play a vital role in organizing and producing persuasive writing, based on the norms and expectations of people involved" (Amiryousefi & Eslami Rasekh, 2010, p. 159). In Adel's (2006) words, "metadiscourse is discourse about evolving discourse or the writer's explicit commentary on her own ongoing text" (p. 2).

As metadiscourse features are regarded as crucially important features for establishing the relationship between writer and reader, Vande Kopple (2012) mentions the reasons for this importance. First, metadiscourse features are used to show that language is a complex and intricate phenomenon. The second reason is that by nature some metadiscourse features can have more than one function, depending on the texts in which they are used.

Studies on Metadiscourse

Studies in the field of metadiscourse features vary. Some of the studies are experimental in nature in that they observe the usage of metadiscourse features in enhancing students writing (see, for example, Cheng & Steffensen, 1996; Pérez & Macià, 2002; Simin & Tavangar, 2009; Vahid Dastjerdi & Shirzad, 2010; Tavakoli & Amirian, 2012), speaking ability (see, for example, Kong & Xin, 2009; Ahour & Entezari Maleki1, 2014), listening ability (see, for example, Heshemi & Khodabakhshzade, 2012) and reading comprehension (see, for example, Flowerdew & Tauroza, 1995; Camiciottoli, 2003; Jalilifar & Alipour, 2007; Behnam & Babapour, 2015).

Other types of studies are categorised in the domain of comparative studies in which metadiscourse features are analysed between various genres (see, for example, Crismore, Markkanen & Steffensen, 1993; Abdollahzadeh, 2000; Beigmohammadi, 2003) or between native speakers and non-native speakers of English (see, for example, Abdollahzadeh, 2003, Davoodifard, 2006; Abdollahzadeh, 2007, Koohi & Moojod, 2012; Bonyadi & Samuel, 2012, Tadayyon & Vasheghani Farahani, 2017).

Research Questions

Using corpora in language studies have attracted a growing attention due to the advances made in computer technology (Mukherjee, 2006). In addition, analysing metadiscourse features in academic writing by analysing a large corpus like the current one has not been reported, to the best knowledge of the author. Considering these issues, this research was

innovative in that it exploited a large, representative, annotated and balanced corpus, which contains 6,506,995 words. Spotting these gaps and with due regard to the above-mentioned issues, this research aims at addressing the following questions:

Q1 What is the distributional pattern of interactive metadiscourse features in the corpus?

Q2 What is the distributional pattern of interactional metadiscourse features in the corpus?

Q3 What is the main tendency of the corpus in terms of interactive or interactional meta discourse features?

Methodology

Design of the Research

This research was comparative, quantitative and non-experimental in nature. It was also a corpus- based research. It was comparative as it tried to compare interactive vs. interactional metadiscourse features in the corpus. It was a non-experimental study as it did not have any manipulation. It was, however, a corpus-based study as it used corpus software and a publicly available corpus.

Instrumentation

For the sake of operationally, any research requires some instrumentations by which the feasibility of it can be assured. Therefore, in order to put this research into practice, a number of instrumentations were used which are as follow.

Metadiscourse Taxonomy

There are various models and taxonomies of metadiscourse (see, for example, Crismore, 1989; Vande Kopple, 1985, 2002; Hyland, 2005; Adel, 2006). These taxonomies of metadiscourse features enjoy some amount of commonalities, despite the fact that they are distinct in categorising metadiscourse features. It seems that the differences between these taxonomies are a matter of terminology as there is a great deal of commonality between them (Jalilifar, Hayati & Don, 2019). However, from among these categories, the one selected for the current research was Hyland's model of metadiscourse features, as it is the most recent and the most comprehensive taxonomy of metadiscourse features. It is also the most straightforward and understandable one which does not have the intricacies of other taxonomies (Ghadyani & Tahririan, 2015). In addition, Hyland's model is appropriate as it takes a functional approach to analysing texts; making it more feasible as compared to other classification (Jalilifar, Hayati & Don, 2019).

In Hyland's taxonomy of metadiscourse features, there are two main categories each is sub-divided into five subcategories. The first main category is interactive metadiscourse features, which has five subdivisions as transitions, frame markers, endophoric markers, evidentials and code glosses. The interactive category of metadiscourse features "concerns the writer's awareness of a participating audience and the way he or she seeks to accommodate its possible knowledge, interests, rhetorical expectations and processing abilities" (Hyland, 2005, p. 49). As a matter of fact, the writer structures the text in such a way that it can meet the requirements of the writer and shape the text so that the audience can react in such a way that the author intends. On the other hand, the interactional metadiscourse features "involve readers and open opportunities for them to contribute to the discourse by alerting them to the author's perspective towards both propositional information and readers themselves (ibid, p. 52). In other words, the writer tries to make the speech as explicit as possible and to produce a coherent, well-structured text so that the audience can easily follow the course of the text.

Table 1
A category of metadiscourse features (Hyland, 2005)

Category	Function	Example
Interactive	Help to guide the reader through the text	Resources
Transitions	express relations between main clauses	in addition, but, thus, and
Frame markers	refer to discourse acts, sequence or stages	finally, to conclude, my purpose
Endophoric markers	refer to information in other parts of the text	noted above; see fig
Evidentials	refer to information from other texts	according to X, Z states
Code glasses	elaborate propositional meanings	Namely, e.g., such as, in other words

Category	Function	Example
Interactional	Involve the reader in the text	Resources
Hedges	withhold commitments and open dialogue	might, perhaps, possible, about
Boosters	emphasise certainty or close dialogue	in fact, definitely, it is clear that
Attitude markers	express writer's attitude to 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

Corpus of the Study

Since the current research was a corpus-based inquiry in nature, it was necessary to compile a corpus that could meet the requirements of the study and the necessities of corpus design such as representativeness and balance (Mcenery & Hardi, 2011). As Curzan (2012) puts it, from among the various functions of corpora, there are three main advantages. Corpora can assist the researchers to detect the texts that are worth investigating. Besides, they can also provide systematicity in that they can be used to cataloguing the linguistics patterns. More, they can show new ways of co-occurrence patterns in texts. All these crucially important considerations require a well-established, clear and coherent data collection procedure because "corpus data can provide a wide variety of examples of actual use without undermining the emphasis on genuine texts studied in context" (Chambers, 2011, p. 98).

In order to have a scientifically acceptable corpus, any type of corpus needs to have some specifications based on which the corpus is designed and analysed. Two of the most important factors are balance and representativeness of the corpus. Representativeness is assured when the various subcategories of the corpus are available (Zanettin, 2012). On the other hand, the larger the corpus size, the better and more robust the results will be. Indeed, "quantitative considerations are at the core of corpus-based studies – and that it must contain texts collected with a specific purpose in mind" (Zanettin, 2012, p. 41).

As Mcenery & Hardies (2012) state, for making a good corpus, some issues need to be taken into account by researcher(s). First and most, the corpus has to be in line with the research question(s) of the study. Moreover, the corpus must be homogeneous in nature; meaning that it must be compiled of texts of similar genre. Finally, the corpus (texts) that we select to

analyse and investigate is required to be compiled from various types of the texts. Considering these issues, this research enjoyed the British Academic Written English Corpus as the data-gathering source. This corpus was designed as a collaboration between the Universities of Warwick, Reading and Oxford Brookes. This corpus was founded by Economic and Social Research Council in 2007. It contains 6,506,995 annotated words in 2761 pieces of proficient students' English writings with the range of 500 to 5000 words. The L1 of the students was not English; rather it contained some other languages as Chinese, Finnish, Japanese Portuguese and Malaya. The texts of the corpus were selected from four main categories of Arts and Humanities, Social Sciences, Life Sciences and Physical Sciences. The corpus is available free of charge to researchers.

Table 2
Information on British Academic Written English Corpus

Token	Words	Sentences	Lemma
8,336,262	6,968,089	293,113	137,598

Table 2 illustrates the initial information on the British Academic Written English Corpus. As can be seen, the corpus is composed of 8,336,262 tokens. It also contains 6,968,089 words and 293,113 sentences. The corpus of the research also contains 137,598 lemma.

Sketch Engine Corpus Software

Any corpus-driven research requires, undoubtedly, computer software (Tymoczko, 1998). Indeed, as far as the text analysis in large quantity is concerned, without using the software, it is impossible and irrational to delve into the corpus. For this purpose, from among the various software, the Sketch Engine was used. Sketch Engine is a Windows supported corpus software which, since its advent in 2003, has been extensively used in different projects such as dictionary compiling, phraseology, collocation studies and text analysis. This programme was designed by Lexical Computing Ltd. (https://www.sketchengine.co.uk/). Apart from being user-friendly, Sketch Engine gives researchers the opportunity to have access to a wide range of raw data from various corpora and languages like National British Corpus, Early English Books Online, and English Web 2013 (McGillivray & Kilgarrif, 2013).

R Statistical Program

Usually, for the statistical analysis in the field of Applied Linguistics, the SPSS software is used. However, this program is subject to some limitations that can question the results. For this purpose, the R software was used in this research. As a matter of fact, R is a freely downloadable software environment for statistical computing and analysis. It can compile and run on a wide variety of UNIX platforms, Windows, and MacOS.

Procedure

In order to put this research into practice, it was necessary to look for all the instances of metadiscourse in the corpus. Since the range of metadiscourse instances is wide, the corpus was analysed by hand to, at first, detect the types (the metadiscourse features); second, by using the software, the tokens (each stance of metadiscourse feature) of each type were searched. This procedure was separately done in both subcategories of the taxonomy, i.e., interactive metadiscourse features and interactional metadiscourse features. Then, the statistical analysis was performed to draw the quantitative conclusions.

Examples of Metadiscourse Features

In order to verify the results of the study, some examples of both categories are represented.

Interactive Metadiscourse Features

Transitions

Fig. 1
Instances of transitions in the corpus

greater levels in the leaves than in the roots.	In addition	, although we did not study the expression of
may occur i.e. in light-grown conditions and	In addition	modifies the expression of this enzyme such
of the gene (only a UV lamp is needed).	In addition	the expression of GFP is easily observed under
equipment is needed to record these findings.	In addition	the red fluorescence of chlorophyll may
mutant plants can be used for further research.	In addition	the equipment used to measure luciferase
fetal malformations) are generated, and	In addition	possibly develop new pre-natal screens

Frame markers

Fig. 2
Instances of frame markers in the corpus

text#178	Hitler's laid-back attitude to work. However,	to conclude	that Hitler was a weak dictator and that this in
text#277	act, leading them to retain it for the future and	to conclude	that this aggressive behaviour is acceptable,
text#278	Previous research has led some psychologists	to conclude	that inferences are drawn through the means of

Endophoric Markers

Fig. 3
Instances of endophoric markers in the corpus

text#178	Company in India. The massive boost to trade,	as noted	by Figures 1 and 2, resulted in the formation of a
text#277	to cooperate, has been largely ineffective,	as noted	by Beigbeder: "The Council has failed to ensure
text#278	, 3 of which only took between 20 and 22 weeks;	as noted	in "UK Merger Control - A Year's Experience of

Evidentials

Fig. 4
Instances of evidentials in the corpus

text#1	ways in which people apply concepts. However	according to	the radical approach there are more
text#1	official statistics do have some uses.	according to	posivitists as long as the statistics are
text#1	to be honest in a self report study. Overall,	according to	the posivitists the main limitations of
text#5	which 'their [women's] sanity was often judged	according to	their compliance with middle-class standards

Code Glosses

text#160	found getting closer to the mean of the means.	In other words	, the range of fluctuate population sizes gets
text#168	by time, but by the ruling elite in any era.	In other words	, memorials can be utilised to either glorify or
text#171	in which commerce played a significant role,	in other words	the development of an infrastructure of
text#172	in favour of mercy in a particular case'.	In other words	not only were women central in the obtaining of

Fig. 5
Instances of code glosses in the corpus

Interactional Metadiscourse Features

Hedges

more reactive to other customers. They seemed,	perhaps	, less accustomed to the unwritten rules of the '
that surrounded them (McCracken 1990:24).	perhaps	the Great Exhibition of 1851 may be said to have
new consumerist middle class model. Except,	perhaps	, for some moral criticism, which was labelled
's relative freedom of action. It is therefore	perhaps	surprising that Rousseau leans entirely on the
as the avant-garde as 'infantile disorders'.	perhaps	due to some drastic changes in Borges' life over

Fig. 6
Instances of hedges in the corpus

Boosters

1939, The National Archives, AIR 12/194	It is clear that	, from January 1939, airframe manufacturers
being mixed together. Results From the results	it is clear that	the bacteria could grow earlier on some plates
, pro, hisDiscussion From the table of results	it is clear that	the genes allowing growth to occur by
in relation to the amount of pNP present.	It is clear that	as the concentration of phosphate is increased
the temperature on enzyme activity From the graph	it is clear that	the temperature at which alkaline phosphatase

Fig. 7
Instances of boosters in the corpus

Attitude Markers

secretaries locked up in rooms for weeks until	they agreed that	they would make the targets set in front of them.
have proposed that leaders are superfluous.	they agree that	providing task structure and showing support
olds participating in education and training,	they agree that	this vision seems a long way off (CEC Brussels

Fig. 8
Instances of attitude markers in the corpus

Fig. 9
Instances of self-mentions in the corpus

Self-mentions

typifications", 'typifications of	our	common-sense typifications which order the
frequently engaged in making connections in	our	descriptions' (Coulter 1979: 15). In this
structural coherence and predictability in	our	perception of the world that we experience. A
the existential reminder of the forces underlying	our	everyday commonsense social interaction. . </p

Fig. 10
Instances of engagement narkers in the corpus

Engagement Markers

the chart below for 'Access to internet'.	You can see that	retailing through the internet may not be the
they believe this. Overall from these results	you can see that	children are no better off health wise with
graph gradient From my results in table 1	you can see that	as you increase , the initial amplitude of the
slightly between figure 3 and figure 4 and	you can see that	*the average gradient of figure 4 is greater than*

Data Analysis

Table 3 demonstrates statistics for distribution of the interactive metadiscourse features in the corpus. As can be seen, the median for the interactive metadiscourse features is 66,823. The mode is 103,110 and the minimum is 9,832. The maximum is 103,110.

Table 3
Statistics for interactive metadiscourse features

N	Valid	208,444
	Missing	0
Median		66,823.00
Mode		103,110
Minimum		9,832
Maximum		103,110

Fig. 11
The distribution of metadiscourse features in the corpus

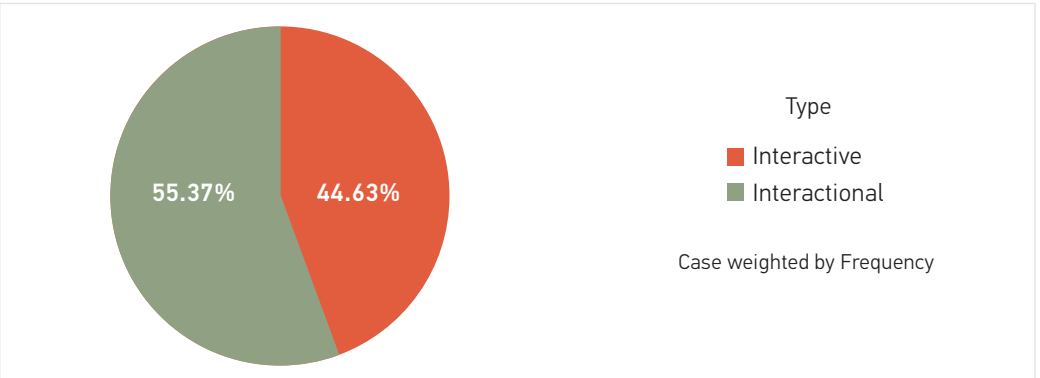

Fig. 11 also represents data of the distributional pattern of the interactive and interactional metadiscourse features in the corpus of the study.

Table 4

Tests of normality of the interactional metadiscourse features in the corpus

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
Frequency	0.296	208,444	0.000

a - Lilliefors significance correction

Table 4 depicts the results of the normality of the interactive metadiscourse features data in the corpus. As can be understood from data, the P value is < 0.05; as a result, the data were not normally distributed as far as the interactive metadiscourse features are concerned. Therefore, non-parametric statistics were applied.

Table 5

Statistics for interactive metadiscourse features

N	Valid	258,630
	Missing	0
Median		104,973.00
Mode		110,049
Minimum		1,966
Maximum		110,049

Table 5 gives information on the distribution of interactional metadiscourse features in the corpus. It is understood from the data that the median is 104,973 and the mode is 110,049. In addition, the minimum and maximum are 1,966 and 110,049, respectively.

Table 6

Tests of normality of the interactional metadiscourse features in the corpus

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
Frequency	0.296	208,444	0.000

a - Lilliefors significance correction

Table 6 provides information on the distribution of interactional metadiscourse features in the corpus. As the results of the Kolmogorov-Smirnov test can show, the level of significance is below < 0.05; as a result, the data were not normally distributed, ensuring the application of non-parametric statistics.

Table 7

The distributional pattern of interactive metadiscourse features in the corpus

			Name					Total
			Code Glosses	Endophoric Markers	Evidentials	Frame markers	Transitions	
Type	Interactive	Count	18,668	66,823	10,011	9,832	103,110	208,444
		% of Type	9.0%	32.1%	4.8%	4.7%	49.5%	100.0%
Total		Count	18,668	66,823	10,011	9,832	103,110	208,444
		% of Type	9.0%	32.1%	4.8%	4.7%	49.5%	100.0%

Table 7 explains the distributional pattern of interactive metadiscourse features in the corpus. As can be understood from data, from among the interactive metadiscourse features, the transitions were the most applicable one by 10,110 counts of the total (49.5%). Transitions were followed by endophoric markers, which constituted 32.1% (66,823) of the total corpus. In the third stand, there were code glosses, 9.0% of the total corpus. With 4.80% and 4.72% of the total corpus, evidentials and frame markers were the least used interactive metadiscourse features.

Fig. 12

The distribution of the interactive metadiscourse features in the corpus

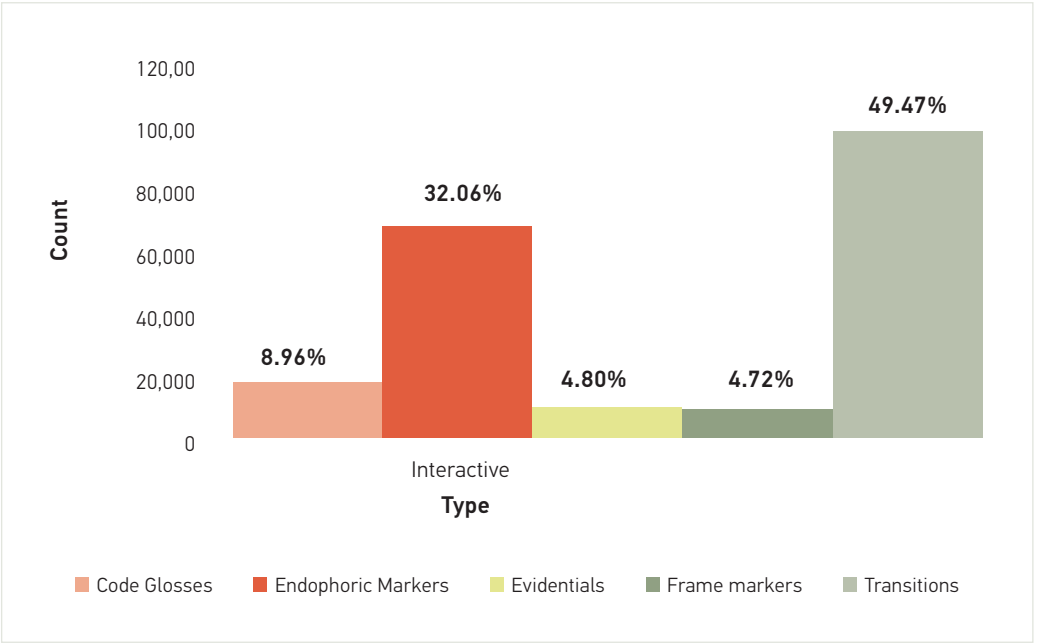

Table 8

The distributional pattern of interactional metadiscourse features in the corpus

			Attitude Markers	Boosters	Engagement Markers	Hedges	Self-mentions	Total
Type	Interactional	Count	9,987	31,655	1,966	110,049	104,973	258,630
		% of Type	3.86%	12.24%	0.76%	42.55%	40.59%	100.0%
Total		Count	9,987	31,655	1,966	110,049	104,973	258,630
		% of Type	3.9%	12.2%	0.8%	42.6%	40.6%	100.0%

Table 8 manifests the pattern of the distribution of interactional metadiscourse features in the corpus of the study. As can be understood from the data, the most frequent interactional metadiscourse features were hedges with 42.55% (110,049). After that, self-mentions with 104,973 (40.59%) were the most prevalent interactional metadiscourse features. With 31,655 counts (12.24%), boosters were found to be the third most used interactional metadiscourse features. Attitude markers constituted 3.86% (9,987) of the total interactional metadiscourse features in the corpus, which were the fourth most often used. The least used interactional metadiscourse features were engagement markers, which constituted 0.76% (1,966) of the total corpus.

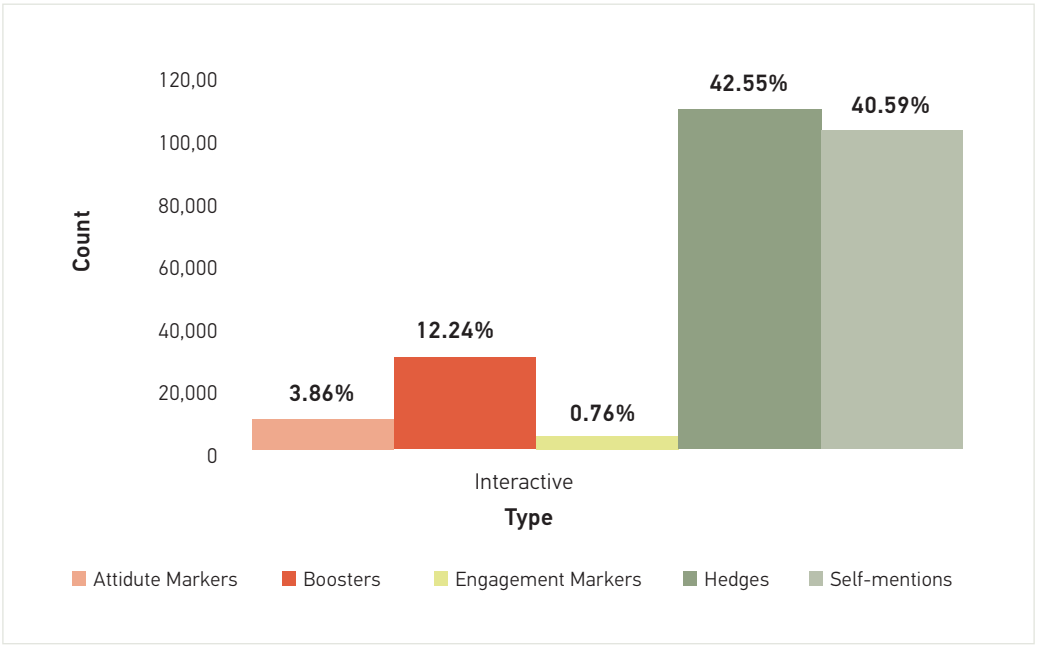

Fig. 13
The distribution of the interactional metadiscourse features in the corpus

Discussion and Results

Academic writing research has gained the attention of the researchers (Letsoela, 2014). As far as writing and metadiscourse features are concerned, previous studies have highlighted the salient role of metadiscourse features in academic writing as a way of constructing interaction between writer and readers (Tse & Hyland, 2006). This research sought to analyse metadiscourse features in a large and balanced corpus of British Academic Written English. For this purpose, the corpus was analysed in terms of metadiscourse features. The Sketch Engine was exploited to extract the metadiscourse features.

The first research question of the study dealt with the distribution pattern of interactive metadiscourse features in the corpus. As the data in table 7 can show, transitions were the most frequent interactive metadiscourse features in the corpus, followed by endophoric markers and code glosses as the second and third most prevalent interactive metadiscourse features. Evidentials with 4.8% of the total interactive metadiscourse features were the fourth used and the least used interactive metadiscourse features were frame markers with only 4.7%.

Transitions are used to connect the main clauses between the sentences (Hyland, 2005). The prominent usage of transitions can add support to the idea that they are an integral part of academic writing. By applying the transitions, authors represent their willingness to produce the text(s) in such a way that readers can unfold their logic. Endophoric markers are used to refer to other parts of the texts. The second stand of endophoric markers can reveal the fact that authors tried to refer to illustrations, examples, sections, parts, and arguments in other parts of the texts that are a feature of academic discourse (Hyland, 2002). Paraphrasing and elaborating the intended meaning is a feature of the academic context. Authors use these techniques to reflect their predictions on reader's knowledge of a proposition (Hyland, 2004). The fact that code glosses were used less often than transitions and endophoric markers can uncover the fact that authors made some efforts to supply extra information to express their indented meanings in a clearer way.

Evidentials can show the fact that writers refer and cite others to justify their intended meaning and persuade the readership. It can be understood that, in the corpus, authors used fewer of these features which can be due to the fact that they were convinced of their propositions

and did not much have to cite others to justify their argumentation. Frame markers that were the least used applicable interactive metadiscourse features can add support to this idea that writers of academic texts were not signalling text boundaries or sequencing different parts of the discourse (Hyland, 2005).

The second research question was concerned with the distribution pattern of interactional metadiscourse features in the corpus. For this end, the instances of interactional metadiscourse features were extracted in the corpus. As the data in Table 8 can show, from among the interactional metadiscourse features, hedges with 42.55% of the whole corpus were the most predominant interactional metadiscourse features. In the second place, there was self-mention with 40.59% of the whole corpus. In the third stance, there were boosters followed by attitude markers and engagement markers, respectively.

That hedges were used as the most significant interactional metadiscourse features can unearth the fact that authors tried to distinguish facts from opinions. Indeed, in academic contexts, facts are presented cautiously and writers assess their propositions and claims in such a way that there is room for alternative voices (Hyland, 2005). In addition, hedges signal the fact that claims were to some extent relative and lacked 100% certainty. As a matter of fact, the extensive usage of hedges revealed that the statements or the propositions offered by authors in the text were mostly based on their own interpretation rather than some certain amount of knowledge.

Self-mentions were the second most applicable interactional metadiscourse features. Self-mentions are among the common features by which the authors put forward their own claims and ideas. The application of self-mentions can lend support to the idea that authors were representing scholarly identity through the interaction with their audience (Hyland, 2001). In other words, the authors showed their strong presence in the text by using the self-mentions.

Boosters as the third most used interactional metadiscourse features can demonstrate the fact that authors were less certain in their propositional and less interested in closing down the argumentation. Indeed, the imbalance of hedges and boosters can depict the fact to what extent authors intend to propose alternative argumentations in the discourse. The low usage of boosters than hedges (uncertainty elements) can reveal the fact that authors were not certain in most of their ideas, propositions, claims, and argumentations, eschewing them to express their 100% certainty in the context.

Attitude markers were regarded as the second least used interactional metadiscourse features. These features are used by authors to show their attitudes towards an argumentation. Indeed, they are used to show effective, not a scientific and logical attitude of authors (Hyland, 2005). In academic discourse, there is less space for using effective features as these features are mostly used in not academic discourse like literature and poetry. As a result, authors of the corpus were not willing to step out of the scientific boundaries.

Finally, the least used interactional metadiscourse features were found to be engagement markers. Usually, engagement markers are used to explicitly attract the attention of the reader towards a proposition (Hyland, 1998). On the other hand, the scientific texts are not the kind of register in which direct instructions on readers can be found. By using these features in a rare mode, this idea can be put that the authors underestimated the presence of the readers during the interaction.

The last research question was about the main tendency of the corpus in terms of interactive or interactional metadiscourse features. For this purpose, comparative statistics were applied. As the data in Fig. 12 can show, from among both interactive and interactional meta-

discourse features, 52.37% of the corpus was dedicated to interactive metadiscourse features; whereas only 44.63% were given to interactional.

Greater reliance of authors on the interactive category of metadiscourse features can deposit that guiding readers in the course of the reading process through discourse organisation and explaining the propositional meanings was very important (Hyland, 2005). In other words, by applying interactive metadiscourse features, authors showed their concern in organising the discourse in such a structured way that the reader can follow the flow of the propositions.

Conclusion

The results of this study showed that interactive metadiscourse features were more prevalent as compared to interactional ones. This prevalence of interactive metadiscourse features can show that at least in the domain of academic writing as a specific genre writers tried to keep their readers fully informed about what they write and how they write as well as representing a clear intention on what the audience can expect to know in the course of interaction. This role was played specifically by transitions and endophoric markers which were found to be the most prevalent metadiscourse features of the interactive category. Moreover, being more interactive in nature as compared with interactional reveals that producing a well-organized and coherent text/ writing was the main concern of the authors. This concern is highlighted when we pay attention to this fact that academic writing, by nature, requires more attention on such concepts as coherence and cohesion, which are mostly attained by metadiscourse features (Littlewood,1966; Bhatia, 2006; Nasiri, 2011).

The findings of this research are in line with those of the Bal-Gezegin (2016) who found out the prevalence of interpersonal metadiscourse features in the English corpus as compared with that of the Turkish corpus. Moreover, this research was in line with that of the Ghanooni and Oghbatalab (2012) who investigated the use of metadiscourse features in academic writing of research articles. In addition, the findings of this research was similar to that of Lin (2005) whose study revealed that students used textual features (68.63%) more than interpersonal features (31.37%). In addition, the results of this research are close to those of Ghafar Samar and Amini (2015) who investigated the pattern of personal and impersonal metadiscourse in academic writing. They showed that interpersonal metadiscourse features were more frequent than textual metadiscourse features.

This study can have different implications for researchers. One implication is for researchers interested in doing corpus linguistics. The procedures taken in the current research can be a step-by-step guide on how to run research in this area of inquiry. In addition, researchers interested in academic written discourse can benefit from the results of this research in that they can understand how metadiscourse features are used in academic discourse and/or writing. Moreover, the results can be useful for researchers in the field of text analysis to understand how metadiscourse features function in written and academic contexts.

This research, like any other one, had some limitations. One limitation was that it was likely that some metadiscourse features, of both kinds, were neglected during the detection process. The other limitation was that this research was focused on academic genre and did not focus on other types of genre or registers. The other limitation was that the current research used Hyland's taxonomy of metadiscourse features as the theoretical framework and did not privilege other categories.

The research at hand can spark off new studies. One study can be done on the distribution of metadiscourse features in spoken discourse, as this area is less researched than the written mode. Indeed, there are some commercially available corpora like British Academic Spoken Corpus for such a research. In addition, researchers can compare the data of British Aca-

demic Written English Corpus with other types of corpora like Cambridge Academic English Corpus, which is commercially available to see how metadiscourse features are used and distributed in these corpora. Furthermore, it is suggested that other researchers embark on doing corpus-driven researches with the focus of comparing metadiscourse features in written corpus vs. spoken corpus like comparing British Academic Written English Corpus and British Academic Spoken Corpus with the aim of comparing metadiscourse features in written and spoken discourse. The last but not the least suggestion is that follow-up studies can be done with the focus on analysing metadiscourse features in academic texts written by native vs. non-native speakers of English as well as other genre.

References

1 Abdi, R., 2000. Interpersonal Metadiscourse Markers in Social Science and Natural Science Research Article Discussion Sections. Unpublished Master's Thesis, University of Tehran, Tehran.

2 Abdollahzadeh, E., 2001. Native and Non-native Writers' Use of Textual Metadiscourse in EL Papers. Unpublished Master's Thesis, University of Tehran, Tehran.

3 Abdollahzadeh, E., 2003. Interpersonal Metadiscourse in ELT Papers by Iranian and Anglo-American Academic Writers. Paper Presented at the International Conference on Multiculturalism in ELT practice at Baskent University, Turkey.

4 Abdollahzadeh, E., 2007. Writer is Presence in Persian and English Newspaper Editorials. Paper Presented at the International Conference on Systemic Functional Linguistics in Odense, Denmark.

5 Ädel, A., 2006. Metadiscourse in L1 and L2 English. Amsterdam: John Benjamins. https://doi.org/10.1075/scl.24

6 Ahour, T., & Entezari Maleki, S., 2014. The Effect of Metadiscourse Instruction on Iranian EFL learners' Speaking Ability. In: English Language Teaching, 7(10), pp. 69–75. https://doi.org/10.5539/elt.v7n10p69

7 Amiryousefi, M., & Eslami Rasekh, A., 2010. Metadiscourse: Definitions, Issues and Its Implications for English Teachers. In: English Language Teaching, 3(4), pp. 159–167. https://doi.org/10.5539/elt.v3n4p159

8 Bhatia, V. K., 2006. Analysing Genre: Some Conceptual Issues. In: Academic Writing in Context: Implications and Applications, (ed.) Hewings, M., pp. 79–92. London: Continuum.

9 Bailey, S., 2003. Academic Writing: A Handbook for International Students. London: Routledge.

10 Bailey, S., 2015. Academic Writing: A Handbook for International Students. London: Routledge.

11 Bal-Gezegin, B., 2016. A Corpus-based Investigation of Metadiscourse in Academic Book Reviews. Procedia – Social and Behavioral Sciences, 232, pp. 713–718. https://doi.org/10.1016/j.sbspro.2016.10.097

12 Behnam, B., & Babapour, M., 2015. EFL learners' Reading Comprehension Ability in the Light of Metadiscourse Awareness. In: Theory and Practice in Language Studies, 5(9), pp. 1850–1856. https://doi.org/10.17507/tpls.0509.12

13 Beigmohammadi, A., 2003. An Investigation into the Patterns of Use of Discourse Features of Intensity Markers in Academic Research Articles of Hard Science, Social Science, and Unpublished Master's Thesis, University of Tehran, Tehran.

14 Biber, D., & Gray, B., 2016. Grammatical Complexity in Academic English: Linguistic Change in Writing. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511920776

15 Boggel, S., 2009. Metadiscourse in Middle English and Early Modern English Religious Texts: A Corpus-based Study. Frankfurt am Main: Peter Lang.

16 Bonyadi, A., & Samuel, M., 2013. Headlines in Newspaper Editorials. SAGE Open, 3(2), pp. 1–7. https://doi.org/10.1177/2158244013494863

17 Bowker, N., 2007. Academic Writing: A Guide to Tertiary Level Writing. Palmerston North: Massey University.

18 Camiciottoli, B. C., 2003. Metadiscourse and ESP Reading Comprehension: An Exploratory Study. In: Reading in a Foreign Language, 15(1), pp. 1–16. Retrieved from https://nflrc.hawaii.edu/rfl/April2003/camiciottoli/camiciottoli.pdf.

19 Chambers, A., 2010. The Learner Corpus as a Pedagogic Corpus. The Cambridge Handbook of Learner Corpus Research, pp. 445–464. https://doi.org/10.1017/CBO9781139649414.020

20 Cheng, X., & Steffensen, M. S., 1996. Metadiscourse: A Technique for Improving Student Writing. In: Research in the Teaching of English, 30(2), pp. 194–181.

21 Cho, K. & Schunn, C. D., 2007. Scaffolded Writing and Rewriting in the Discipline: A Web-based Reciprocal Peer Review System. In: Computers & Education, 48(3), pp. 409–426. https://doi.org/10.1016/j.compedu.2005.02.004

22 Crismore, A., 1989. Talking with Readers: Metadiscourse as Rhetorical Act. New York: Peter Lang Publishers.

23 Crismore, A., Markkanen, R., & Steffensen, M. S., 1993. Metadiscourse in Persuasive Writing: Study of Texts Written by American and Finnish University Students. In: Written Communication, 10(1), pp. 39–71. https://doi.org/10.1177/0741088393010001002

24 Curza, A. 2012. The Electronic Life of Texts: Insights from Corpus Linguistics for all Fields of English. In: English Corpus Linguistics: Crossing Paths, (ed.) Kytö, M., pp. 7–21. Amsterdam: Rodopi. https://doi.org/10.1163/9789401207935_003

25 Davoodifard, M., 2006. A Contrastive Analysis of Hedging in English and Persian Research Articles: Linguistic and Cultural Variations Across Languages and Disciplines. An Unpublished Thesis, University of Esfahan, Iran.

26 Defazio, J., Jones, J., Tennant, F., & Hook, S. A., 2010. Academic Literacy: The Importance and Impact of Writing Across the Curriculum – A Case Study. In: Journal of the Scholarship of Teaching and Learning, 10(2), pp. 34–47.

27 Farahani, M. V., & Sbetifard, M., 2017. Metadiscourse Features in English News Writing Among English Native and Iranian Writers: A Comparative Corpus-based Inquiry. In: Theory and Practice in Language Studies, 7(12), 1249. https://doi.org/10.17507/tpls.0712.12

28 Flowerdew, J., & Tauroza, S., 1995. The Effect of Discourse Markers of Second Language Lecture Comprehension. In: Studies in Second Language Acquisition, 17, pp. 435–458. https://doi.org/10.1017/S0272263100014406

29 Ghadyani, F., & Tahririan, M. H., 2015. Interactive Markers in Medical Research Articles Written by Iranian and Native Authors of ISI and Non-ISI medical journals: A Contrastive Metadiscourse Analysis of Method Section. In: Theory and Practice in Language Study, 5(2), pp. 309–317. https://doi.org/10.17507/tpls.0502.10

30 Ghafar Samar, R., & Amini, S. H., (n.d.). A Comparative Study of Personal and Impersonal Metadiscourse in Academic Writing. In: Journal of Studies in Learning and Teaching English, 4(2), pp. 53–66.

31 Ghafoori, N., & Oghbatalab, R., 2012. A Comparative Study of Metadiscourse in Academic Writing: Male vs. Female Authors of Research Articles in Applied Linguistics. In: The Journal of Applied Linguistics, 5(1), pp. 88–113.

32 Gholam, M., Tajalli, J., & Shokrpour, N., 2014. An Investigation of Metadiscourse Markers in English Medical Texts and Their Persian Translation: Based on Hyland's Model. In: European Journal of English Language and Literature Studies, 2(2), pp. 1–41.

33 Heng, C. S., & Tan, H., 2010. Extracting and Comparing the Intricacies of Metadiscourse of Two Written Persuasive Corpora. In: International Journal of Education and Development using Information and Communication Technology, 6(3), pp. 124–146.

34 Hyland, K., 1998. Persuasion and Context: The Pragmatics of Academic Metadiscourse. In: Journal of Pragmatics, 30(4), pp. 437–455. https://doi.org/10.1016/S0378-2166(98)00009-5

35 Hyland, K., 2000 Disciplinary Discourses: Social Interactions in Academic Writing. London: Longman.

36 Hyland, K., 2001. Bringing in the Reader: Addressee Feature in Academic Articles. In: Written Communication, 18(4), pp. 549–574. https://doi.org/10.1177/0741088301018004005

37 Hyland, K. 2002. Authority and Invisibility: Authorial Identity in Academic Writing. In: Journal of Pragmatics, 34(8), pp. 1091-1112. https://doi.org/10.1016/S0378-2166(02)00035-8

38 Hyland, K., & Tse, P., 2004. Metadiscourse in Academic Writing: A Reappraisal. In Applied Linguistics, 25(2), pp. 156–177. https://doi.org/10.1093/applin/25.2.156

39 Hyland, K., 2005. Metadiscourse: Exploring Interaction in Writing. London: Continuum.

40 Hyland, K., & Jiang, F., 2018. In this Paper We Suggest: Changing Patterns of Disciplinary

Metadiscourse. English for Specific Purposes, 51, pp. 18–30. https://doi.org/10.1016/j.esp.2018.02.001

41 Hyland, K., 2019. Metadiscourse: Exploring Interaction in Writing. London: Continuum.

42 Jalilifar, A., & Alipour, M., 2007. How Explicit Instruction Makes a Difference: Metadiscourse Markers and EFL Learners' Reading Comprehension Skill. In: Journal of College Reading and Learning, 38(1), pp. 127–148. https://doi.org/10.1080/10790195.2007.10850203

43 Jalilifar, A., Hayati, S., & Don, A., 2018. Investigating Metadiscourse Markers in Book Reviews and Blurbs: A Study of Interested and Disinterested Genres. In: Studies About Languages, 33, pp. 90–107. https://doi.org/10.5755/j01.sal.0.33.19415.

44 Kong, R., & Xin, X., 2009. Empirical Study on Metadiscourse in Chinese EFL Learners' Oral Communication. In: CELEA Journal, 32(1), pp. 52–64.

45 Kopple, W. J., 1985. Some Exploratory Discourse on Metadiscourse. In: College Composition and Communication, 36(1), pp. 82–93. https://doi.org/10.2307/357609

46 Kuhi, D., & Mojood, M., 2014. Metadiscourse in Newspaper Genre: A Cross-linguistic Study of English and Persian Editorials. In: Procedia – Social and Behavioral Sciences, 5(1), pp. 1046–1055. https://doi.org/10.1016/j.sbspro.2014.03.515

47 Lautamatti, L., 1978. Observations on the Development on the topic of in Simplified Discourse. In: Text linguistics, Cognitive Learning and Language Teaching, (eds.) Kohonen, V., & Enkvist, N. E. Turku: University of Turku.

48 Letsoela, P. M., 2014. Interacting with Readers: Metadiscourse Features in National University of Lesotho Undergraduate Students' Academic Writing. In: International Journal of Linguistics, 5(6), pp. 138–153. https://doi.org/10.5296/ijl.v5i6.4012

49 Lin, C. Y., 2005. Metadiscourse in Academic Writing: An Investigation of Graduate Students' MA Thesis in Taiwan. In: Taiwan Journal of TESOL, 2(1), pp. 1–66.

50 Littlewood, W., 1996. Academic Writing in Intercultural Contexts: Integrating Conventions and Personal Voice. In: Hong Kong Journal of Applied Linguistics, 1, pp. 1–18.

51 McEnery, T., & Hardie, A., 2011. Corpus Linguistics: Method, Theory and Practice. Cambridge, England: Cambridge University Press. https://doi.org/10.1017/CBO9780511981395

52 McGillivray, B., & Kilgrriff, A., 2013. Tools for Historical Corpus Research and a Corpus of Latin. In: New Methods in Historical Corpora, (eds.) Bennett, P., Durrell, M., Scheible, S., & Whitt, R. J. Tübingen: Narr.

53 Mukherjee, J., 2006. Corpus Linguistics and Language Pedagogy: The State of the Art- and Beyond. In: Corpus Technology and Language Technology: New Resources, New Tools, New Methods, pp. 5–24. Wien: Peter Lang.

54 Nasiri, S., 2011. A Contrastive Study of Hedges in Environmental Sciences RAs. Unpublished Master's thesis. Urmia: IAU, Science and Research Branch, Department of English.

55 Pérez, M. A., & Macià, E. A., 2002. Metadiscourse in Lecture Comprehension: Does it Really Help Foreign Language Learners? In: Atlantis Journal, 24(1), pp. 3–21.

56 Simin, Sh., & Tavangar, M., 2009. Metadiscourse Knowledge and Use in Iranian EFL Writing. In: Asian EFL Journal, 11(1), pp. 230–255.

57 Swales, J., 1990. Genre Analysis: English in Academic and Research Settings. Cambridge: Cambridge University Press.

58 Vahid Dastjerdi, H., & Shirzad, M., 2010. The Impact of Explicit Instruction of Metadiscourse Markers on EFL Learners' Writing Performance. In: The Journal of Teaching Language Skills, 2(2), pp. 155–174.

59 Vande Kopple, W. J., 2002. Metadiscourse, Discourse, and Issues in Composition and Rhetoric. In: Discourse Studies in Composition, (eds.) Barton, F., & Stygall, C., Cresskill: Hampton Press. pp. 91–113.

60 McEnery, T., & Hardie, A., 2012. Corpus Linguistics: Method, Theory and Practice. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511981395

61 Tadayyon, M., & Vasheghani Farahani, M., 2017. Exploring Discourse Markers Used in Academic Papers: A Comparative Corpus-based Inquiry of Iranian and English Native Writers. In: Iranian EFL Journal, 13(2), pp. 130–150.

62 Tse, P., & Hyland, K., 2006b. Gender and Discipline: Exploring Metadiscourse Variation in Academic Book Reviews. In: Academic Discourse Across Disciplines, (eds.) Hyland, K., & M., Bondi, pp. 177–202. Bern: Peter Lang.

63 Tymoczko, M., 1998. Computerized Corpora and the Future of Translation Studies. In: Meta: Journal des traducteurs, 43(4), pp. 652–660. https://doi.org/10.7202/004515ar

64 Vande Kopple, V. J., 2012. The Importance of Studying Metadiscourse. In: Applied Research in English, 1(2), pp. 37–44.

65 Weisser, M., 2016. Practical Corpus Linguistics: An Introduction to Corpus-based Language Analysis. Hoboken: John Wiley & Sons. https://doi.org/10.1002/9781119180180

66 Zanettin, F., 2016. Translation-driven Corpora: Corpus Resources for Descriptive and Applied Translation Studies. London, England: Routledge: Routledge.

67 Zarei, G., & Mansoori, S., 2007. Metadiscourse in Academic Prose: A Contrastive Analysis of English and Persian Research Articles. In: The Asian ESP Journal, 3(2), pp. 24–40.

Santrauka

Mehrdad Vasheghani Farahani. Metadiskursas akademiniuose anglų kalbos tekstuose: britų akademinės rašytinės anglų kalbos tekstyno tyrimas

Šiame darbe aprašomas tekstynų lingvistikos srities tyrimas, kuris susitelkia ties metadiskurso elementų britų akademinės rašytinės kalbos tekstynuose analize. Dėl šios priežasties šio tyrimo duomenų šaltiniu buvo pasirinktas laisvos prieigos Britų akademinės rašytinės anglų kalbos tekstynas, kuriame yra 6 968 089 žodžiai. Hylando sudaryta metadiskurso elementų taksonomija yra šio tyrimo teorinis pagrindas, o „R“ programa buvo naudojama kaip statistinė programinė įranga statistikos analizei atlikti. Duomenys rodo, kad tekstyne esantys interaktyvūs metadiskurso elementai vyrauja labiau negu interakciniai matadiskurso elementai. Interaktyvių elementų kategorijoje dažniausiai naudojami teksto jungtukai ir endoforiniai žymekliai, o interakcinėje kategorijoje – sąšvelniai ir skatinamieji žodžiai. Tai, jog interaktyvūs metadiskurso elementai yra ryškiai pastebimi tekstyne, patvirtina hipotezę, kad rašytojams svarbiau yra sisteminti diskursą nei palaikyti ryšį su skaitytojais. Šio tyrimo rezultatai gali būti naudingi gretinamosios ir tekstynų lingvistikos tyrėjams.

About the Author

Mehrdad Vasheghani Farahani

Leipzig University, Germany

Research Interests
Corpus linguistics, corpus based translation studies, contrastive analysis, translation theories

Address
Phillip-Rosenthal Street 33, 04103 Leipzig, Germany

E-mail
mehrdadfarahani1365@gmail.com