METADISCOURSE IN ACADEMIC WRITTEN AND SPOKEN ENGLISH: A COMPARATIVE CORPUS-BASED INQUIRY

MEHRDAD VASHEGHANI FARAHANI
University of Leipzig, Germany
mv64nyjo@studserv.uni-leipzig.de

Abstract
This paper reports on a comparative study performed in the field of Corpus Linguistics. The objective of the research was to analyze the distributional pattern of interactive and interactional metadiscourse features in two modes of academic spoken and written English. For this reason, a list of metadiscourse characteristics was gathered. By using the Sketch engine software, all the words were scrutinized in the corpus and their concordance lines were analyzed one by one in both corpora (British Academic Written English Corpus and British Academic Spoken English Corpus). As the data can show, in both corpora, the general propensity of the authors was towards the interactive metadiscourse features. In addition, in the written corpus, the transitions and endophoric markers were used more often; while in the spoken, endophoric markers and transitions were the most frequently applied metadiscourse features. In the interactional metadiscourse features, hedges and self-mentions were the most frequent in the written form; whereas in the spoken, self-mentions and boosters were used moe often.

Keywords: metadiscourse features, comparative linguistics, academic discourse, written and spoken mode

1. Introduction

Writers or speakers do not produce a piece of text or an oral message to simply communicate and/or exchange information; rather, they look for the ways to ensure that the flow of communication has been successful and the reader can understand the propositions, or the message, offered by the authors (Amiryousefi and Eslami Rasekh, 2010). In this regard, writing or speaking, especially in the domain of academic context, is perceived to be a social process between the writers and speakers, on the one hand, and readers and listeners, on the other hand (Hyland and Tse, 2004; Hyland, 2004). From various genres, the academic genre is used for miscellaneous purposes (Bailey, 2003; Ghahremani Mina and Biria,

2017). Indeed, this special kind of genre is used for reporting the results of research, answering research questions, discussing a scientific inquiry and synthesizing research programmes done by others.

2. Academic Language and its features

The important role of articles in academic context cannot be underestimated (Hyland, 2005). Bowker (2007) points out some of the most salient features of academic writing by which it can be distinguished from other types of writings. For her, one major difference is the application of punctuation and grammar, which follows very strict rules. Indeed, academic writing adheres to a very precise usage of punctuation and shuns the haphazard exploitation of punctuation marks. Apart from punctuation and grammar, academic writing focuses, to a great extent, on abstract propositions, which cannot be explained in physical formats. Moreover, academic writing necessitates the application of citations and reference to other works, which is done for producing a more rational proposition. In other words, "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" (Reza Zarei & Mansoori, 2007, p. 25).

Like writing, academic spoken English has a number of features that make it different from academic written English. These features can include such elements as formality of language, explicitness, precision, hedging, responsibility, accuracy, level of complexity, objectivity, planning, and organizing (http://www.uefap.com/speaking/feature/intro.htm). The spoken medium demonstrates that the speaker is observing the flow of information. It is indeed a manifestation of speaker's endeavour related to communication and maintenance. In this regard, "speaking is an interactive process which leads to constructing meaning that involves the development of a particular type of production, reception, and information processing in its typical grammatical, lexical, and discourse patterns" (Ahour & Entezari Maleki, 2014, p. 69).

The general accepted features of spoken mode can include "Immediateness of presence, emergent unanticipated consequences, recognition of strange otherness, collaborative orientation, vulnerability, mutual implication, temporal flow, genuineness and authenticity" (Cissna & Anderson, 2002, pp. 9-11). In addition, features such as stringing together of idea units without connectives or coordinating conjunctions, speakers' reference to him-/her-self (I, we, and us) and likewise – you, explicitness of speaker's mental processes as revealed by such expressions as "I thought… I", "I had no idea how …" ("which are conspicuously absent in written language "(Chafe, 1982, p. 46) are other characteristics of spoken discourse.

As far as production of a message is concerned, spoken and written language differ in a number of ways that have been emphasized by Brown and Yule (1993).

These differences are seen in such variables as the topic of discussion and the role of participants (Hudson, 1980). Spoken language tends to be less formal and precise than written language. In this regard, Brown, and Yule (1993) point out the following differences:

- The syntax of the spoken language is less structured;
- There are lots of metalinguistic features in written language;
- In spoken language, there are fewer premodified noun phrases;
- The informality of speech is prevalent in spoken language;
- Passive voice is more frequent in written language.

One of the salient features in the academic genre (both written & spoken) is the application of metadiscourse features (Latawiec, 2012). The term metadiscourse was first coined by Harris (1959) to refer to the way of perceiving the language in real contexts, which assists the writers or speakers to guide the receivers' understanding of the message (Ghahremani Mina & Biria, 2017). To put it simply, metadiscourse is defined as linguistic resources "used to paganize a discourse or the writer's stance towards either its content or the reader" (Hyland, 2004, p. 157). In other words, metadiscourse is the author's linguistic and rhetorical manifestation in the text and/or his or her presentation in order to "bracket the discourse organization and the expressive implications of what is being said" (Schiffrin, 1980, p. 231). In another definition, metadiscourse is defined as "language in text which talks about the text rather than propositional content" (Thompson, 2003, p. 6). Hyland (1998) puts forward a definition that metadiscourse features are regarded as those aspects of languages by which the author(s) can represent themselves in the text and facilitate the communication between the authors and the readers.

Metadiscourse features are salient features of academic writing as "it represents writers' attempts to present and negotiate propositional information in ways that are meaningful and appropriate to a particular disciplinary community" (Hyland, 2004, p.136) and in an academic presentation, for producing an interactive process between speaker and hearer, speakers deploy metadiscourse features (Latawiec, 2012). As a result, metadiscourse is regarded an essential specification of spoken interactions between speaker and hearer (Penz, Graf and Marko, 2016).

Metadiscourse features are considered to be "a way of understanding how academic writers express their interpersonal understandings, how they shape their propositions to create convincing, coherent discourse in particular social and institutional contexts" (Hyland, 2004, p.138). For this reason, metadiscourse features have been widely used in various aspects of English as an international language. According to Crismor and Abdollahzadeh (2010), the burgeoning research on metadiscourse features is classified into various domains. 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 and Steffensen, 1996; Pérez and Macià, 2002; Vahid Dastjerdi and Shirzad, 2010), speaking ability (see for example Kong and Xin, 2009), listening skills (Heshemi

and Khodabakhshzade, 2012) and reading comprehension (see for example Flowerdew and Tauroza;1995; Camiciottoli, 2003; Jalilifar and Alipour, 2007).

Other types of studies are categorized in the domain of comparative researchers in which metadiscourse features are analyzed across various genres (see for example Crismore, Markkanen & Steffensen, 1993; Beigmohammadi, 2003), or between native speakers and non-native speakers of English (see for example Abdollahzadeh, 2003, Davoodifard, 2006; Abdollahzadeh, 2007).

3. Research questions and methodology

Analyzing metadiscourse features by applying corpus-based methodology is a feasible approach to delve into the changes of languages in various aspects (Boggel, 2008). Indeed, analyzing metadiscourse features in a comparative mode and based on two largely annotated representative corpora is a field of inquiry which, to the best knowledge of the researcher, lacks research, thus making this study a novel inquiry into the nature of the phenomena. Understanding this gap, this research aimed at analyzing metadiscourse features in two large, balanced, representative, and available written and spoken corpora of the academic genre to investigate the distribution pattern of selected features. Accordingly, the following research questions were proposed.

Research Questions
Q1: What is the distributional pattern of interactive metadiscourse features in written corpus?
Q2: What is the distributional pattern of interactional metadiscourse features in written corpus?
Q3: What is the distributional pattern of interactive metadiscourse features in spoken corpus?
Q4: What is the distributional pattern of interactive metadiscourse features in spoken corpus?
Q5: Are there statistically any significant differences between distributive pattern of metadiscourse features in spoken and written corpora?

Methodology
Design of the Research
Regarding the design of the research, it is conspicuous from the title that this research was comparative, non-experimental, synchronic and corpus-based. It was a comparative research in that it was intended to compare the academic genre of two modes, i.e., spoken and written. It was corpus-based as it exploited corpus software to extract the metadiscourse features and to look for the concordance lines in two sub corpora.

Instrumentation
The Corpus of the Research

Weisser (2016) defines 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" (p.13). Corpus Linguistics as a field of study that studies language in large quantity and deals with texts that can be read by the machine (Mukherjee, 2006, Mcenery and Hardie, 2012 and Zanettin, 2012) has gained much attention due to the advances in computer science, making it an appropriate methodology to study a large collections of texts (Mcenery & Hardie, 2012). Since compiling a representative, reliable, and balanced corpus is a time consuming process (Kruger, 2002), two commercially available corpora were exploited in the study, viz. `British Academic Written English Corpus and British Academic Spoken English Corpus.

The British Academic Written English Corpus is compiled using university-level students' writings of the 21st century. The essays of the students range from 500 to 5000 words in length. It is a corpus of 6, 506, 995 words; it is representative and balanced as it covers a wide range of topics. It includes contributions pertaining to arts, humanities, social sciences, life sciences, and physical sciences. The writings are obtained from undergraduate (BA level) and MA level students and their assignments are all annotated. The corpus is available free of charge and can be accessed to via Sketch Engine Software; it is located at https://warwick.ac.uk/fac/soc/al/research/collections/bawe/ (2018).

The British Academic Spoken English Corpus is compiled at the Universities of Warwick and Reading. The corpus consists of 160 lectures and 39 seminars from various departments of the two universities; all contributions are from the years 1998 to 2005. As is typical of spoken corpora, this corpus has been transcribed and tagged. The spoken corpus, parallel to the written one, has a high index of representativeness and is well-balanced as it also covers various topics related to arts and humanities, life and medical sciences, physical sciences, and social studies and Sciences. (cf. http://www.coventry.ac.uk/research/research-directories/current-projects/2015/british-academic-spoken-english-corpus-base/.

Table 1. Information about British Academic Written English Corpus			
Number of Words	Number of Tokens	Number of Lemmas	Number of Sentences
6,968,089	8,336,262	137,598	293,113

Table 2. Information about British Academic Spoken English Corpus

Number of Words	Number of Tokens	Number of Lemmas
1,186,290	1,252,256	24,653

Tables 1 and 2 report on the preliminary information related to the British Academic Written English Corpus and the British Academic Spoken Corpus, respectively. As can be seen, the former corpus (academic written data) consists of 6,968,089, 8,336,262 tokens, 137,598 lemmas, and 293,113 sentences; whereas, the latter (academic spoken data) contains 1,186,290 words, 1,252,256 tokens, and 24,653 lemmas. The main reason why the number of words and tokens in the written corpus is higher than that of the spoken corpus is that usually compiling a spoken corpus is a time consuming and an arduous task. It requires transcription and numerous annotations, which is time consuming and discouraging (Lovei, Dembryii & Hardie, Brezinai & Tony McEnery,2017).

Metadiscourse Taxonomy
Metadiscourse is important in that it helps to understand language in use; representing the efforts made by the author to guide the perception of the text on the part of the receiver (Harris, 1959). In order to explore metadiscourse features, it was necessary to provide a taxonomy of such features. There are various classifications of metadiscourse (see for example Crismore, 1989; Vande Kopple, 1985, 2002; Hyland, 2005; Adel, 2006). For the purposes of the reported research the taxonomy proposed by Hyland (2005) was adopted as this classification is the most transparent and seems more practical than numerous others with complex problematic categories (Ghadyani& Tahririan, 2015).

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). 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, 52)

Table 3. 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	As noted above; see fig.
Evidentials	refer to information from other texts	According to X; Z states
Code glosses	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	Emphasis certainty or close dialogue	in fact; definitely; it is clear that
Attitude markers	Express writer's attitude to the proposition	Unfortunately; I agree; surprisingly
Self-mentions	Explicit reference to the author(s)	I; we; my; me; our
Engagement markers	Explicitly build a relationship with the reader	consider; note; you can see that

Sketch Engine Software

Any corpus-driven research requires, undoubtedly, a reliable computer software for extracting the data and analyzing them (Tymoczko, 19998); especially in the case of large databases, it is quite impossible and irrational to delve into the texts without proper software. There are lots of corpus-based software such as Wordsmith, the Sketch engine, and Lexa. For the reported research programme, 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 magnificent tool was designed by Lexical Computing Ltd. (https://www.sketchengine.co.uk/). Apart from being user-friendly, Sketch Engine gives the researchers the opportunity to have access to a wide range of raw data from various corpora and languages, like the National British Corpus, Early English Books Online, English Web 2013, *inter alia* (McGillivray & Kilgarrif, 2013).

4. Data: presentation and discussion

Examples of Metadiscourse Features in Two Corpora

Concordance lines are among the best tools for extracting and analyzing the language features manifested in natural contexts where they are used (file:///C:/Users/Mehrdad/Downloads/11637.pdf). For the purposes of the reported study, the concordance lines which were supposed to illustrate

metadiscourse features were extracted. In what follows, selected such examples are presented.

Interactive Metadiscourse Features in Written and Spoken Corpora

Table 4: Transitions

the war of Araucanian independence Araucanian	**in other words**	began to become a poetic way of saying Chile for
i mean there it 's really striking similarities	**in other words**	the frontiers i think of these new republics
th as far as anyone has been able to discern at all	**in other words**	mere geographical or economic unity is not
against a single one of you or your brothers	**in other words**	He's here in this very early phase of the move

Table 5: Frame markers

they were [[voiced pause]] eventually forced	**to conclude**	that this business of allophonic marking of
the three- dimensional image in holography so	**to conclude**	what we 've done today is the first thing to remember
are not just blind thing so well i sort of hasten	**to conclude**	but i think this is [[voiced pause]] all i want
other European powers and the European Union	**to conclude**	my lecture a little earlier than i had planned

Table 6: Endophoric markers

inequality [[voiced pause]] but it 's worth	**noting that**	something like seventy per cent of medical
so let me just mention that the next it 's in your	**notes that**	the next is that you sometimes have data at two
destination [[voiced pause]] operands but	**note that**	they are specified by quoting the the address of
of you that are wide awake unlike myself will	**note that**	this S-plus-alpha- all-squared is exactly

Table 7: Evidentials

was mad for instance [[voiced pause]]	**according to**	Suetonius, he was standing beside the great
to send a force there so Berikos ostensibly	**according to**	our sources and it's Dio Cassius who's fairly
you can well understand why the few times	**according to**	Suetonius that he was trundled out in order to
warning by the example so the people who pass by	**according to**	Mr Fairchild take warning from this example and

Table 8: Code glosses

i mean there it 's really striking similarities	**in other words**	the frontiers i think of these new republics
th as far as anyone has been able to discern at all	**in other words**	mere geographical or economic unity is not
against a single one of you or your brothers	**in other words**	He's here in this very early phase of the move
Americans are the people who are n't Spanish so	**in other words**	the leaders of independence deliberately

Interactional

Table 9: Hedges

only assessed the anxiety/arousal levels.	**I am not sure**	whether the construct validity of the
exercising some form of authority over her.	**I am not sure**	if this behavior comes from cultural
their ideas, in which they actively took part.	**I am not sure**	however about the real reasons, why this
his room where he keeps stuffed toys on his bed,	**I am not sure**	why this is - they possibly remind him of his own

Table 10: Boosters

and so valuable progress is being made * * As	**it is clear that**	eyewitness error can occur in a number of
many aspects of the lives of older people. * *	**It is clear that**	selective attention deficits do exist and
(e.g. language, memory and problem solving)	**it is clear that**	attentional deficits affect many aspects of
in selective attention (Shaw, 1991).	**It is clear that**	the most important findings arising from
as well as the Metro. On closer inspection,	**it is clear that**	Centro is dominant in each of the public

Table 11: Attitude markers

highly successful; it provides the reader (and	**hopefully**	the wider academic community) with a fresh
. Stemming from this examination it shall	**hopefully**	be shown that both liberal and socialist
to this century's dominant worldviews will	**hopefully**	reveal their major failings and advantages,
analyzing each of these components will	**hopefully**	yield insight into the use and effectiveness of
with the business will be dealing with a (	**hopefully**	) safer and better-managed organization. *</p>*

Table 12: Self-mentions

NIMBER: 12/12/2005'And therefore, Reader,	**I myself**	am the subject of my book'. `<p>` M. Montaigne, The
for writing these Essays. He declares that, '	**I myself**	am the subject of my book' and at the same time it
on horse and in chariots, had been set up in Rome;	**I myself**	removed them, and with the money I set golden
, for a start I must look at the role that	**I myself**	play in the business, but I also must look at what

Table 13: Engagement markers

, moved out of trade, although it is necessary to	**remember that**	a significant minority of noble families did
yet to be discovered. It is also important to	**remember that**	despite the association with these individual
deformity). However, it is important to	**remember that**	a negative barium enema doesn't definitely
Mr where available. It is however important to	**remember that**	the majority of these investigations (with the
a normed division algebra. • `</p><p>` It is important to	**remember that**	a normed division algebra is not necessarily a
dimension 16, 32, 64,..., 2n. It is important to	**remember that**	these hypercomplex number systems are not

Data Analysis

For the statistical analysis, the SPSS software (version 21) was used. It is worth noting that since the number of words in both corpora was not necessarily equivalent, the frequency of interactive and interactional metadiscourse features was calculated separately for each corpus; then the ratio of them was examined.

The Analysis of interactive and interactional metadiscourse features in the written corpus

For a proper understanding of the distribution pattern of interactive metadiscourse features in the written corpus, their frequency was calculated for each subcategory (table 14).

Table 14: The distribution of interactive metadiscourse features in Written Corpus

	Transition	Frame Markers	Endophoric Markers	Evidentials	Code Glosses
Count	103110	9832	66823	10011	18668
% of Total	22.1%	2.1%	14.3%	2.1%	4.0%
Count	103110	9832	66823	10011	18668
% of Total	22.1%	2.1%	14.3%	2.1%	4.0%

Table 4 represents the distribution pattern of interactive metadiscourse features in the written corpus. As it is depicted, in all interactive metadiscourse

features, transitions are the most frequent ones (22.1%) followed by endophoric markers (14.3%) and code glosses (4.0%), as the second and third most frequent, respectively. The fourth and fifth, least frequently distributed interactive metadiscourse features in the written corpus were evidential and frame markers, both at the level of 2.1%.

In order to track the distribution pattern of interactional metadiscourse features, their frequency was also calculated.

Table 15: The distribution of interactional metadiscourse features in the written corpus

	Hedges	Boosters	Attitude Markers	Self-mentions	Engagement Markers
Count	110049	31655	9988	104973	1966
% of Total	23.6%	6.8%	2.1%	22.5%	0.4%
Count	110049	31655	9988	104973	1966
% of Total	23.6%	6.8%	2.1%	22.5%	0.4%

Table 5 exhibits the way in which interactional metadiscourse features were distributed in the written corpus. As can be seen from the data, from among the interactional metadiscourse features in the written discourse, with 23.6%, hedges were the most frequently used metadiscourse items in the written corpus. After hedges, self-mentions were the second most frequent interactive element (22.5%). The third most frequent interactive metadiscourse features in the written corpus were boosters, which constituted the 6.8% of the total. Attitude markers were the fourth on the scale, with only 2.1%. The least frequently used interactive

Figure 1 the bar chart of the distribution of interactive and interactional
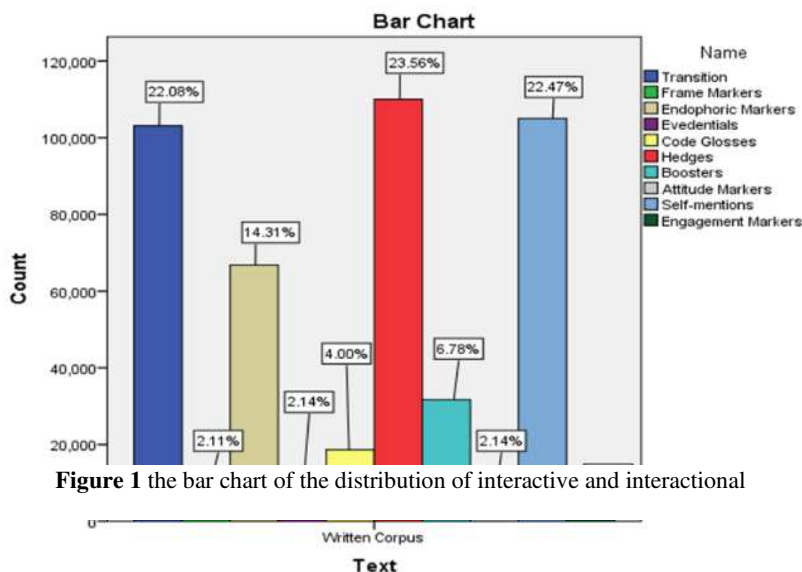
metadiscourse features in the written corpus with only 0.4%, were engagement markers. The distribution has been shown in Figure 1 below.

Table 16: Cross-tabulation of the interactive and interactional metadiscourse features in the written corpus

| | | | Type | | |
			Interactive	Interactional	Total
Text	Written Corpus	Count	208444	258630	467074
		% of Total	44.6%	55.4%	100.0%
Total		Count	208444	258630	467074
		% of Total	44.6%	55.4%	100.0%

Table 6 shows the overall orientation of the interactive and interactional metadiscourse features in the written corpus. A can be seen from the data, the number of interactive metadiscourse features present in the corpus is as high as 208444; whereas the total number of interactional metadiscourse features are 258630; constituting together a number of 467074 counts. It is noteworthy that the number of interactive metadiscourse features identified in the corpus is lower than that of the interactional metadiscourse features in the sense that while interactive metadiscourse features constitute 44.6%, the interactional ones constitute 55.4%, which has been illustrated in Fig. 2 below.
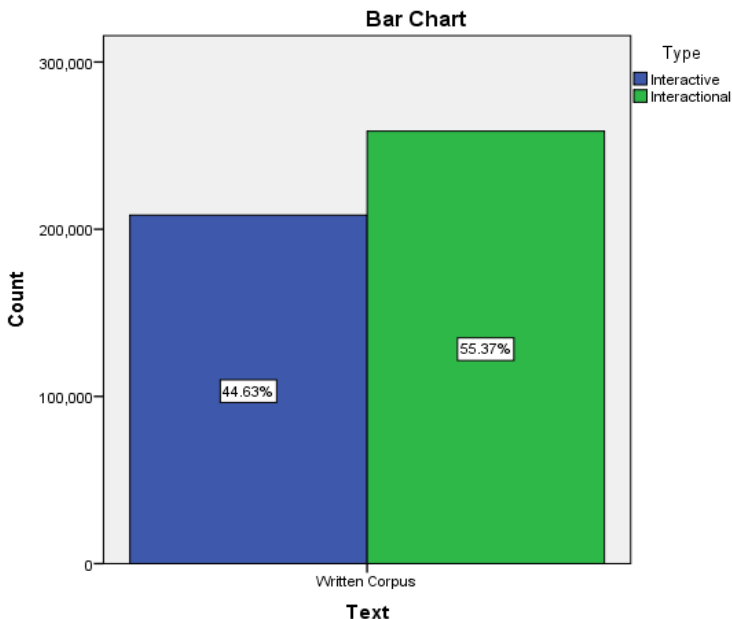

Figure 2. The bar chart on the distribution of interactive and interactional metadiscourse features in the written corpus.

The Analysis of interactive and interactional metadiscourse features in the spoken corpus

As with the written corpus, in order to understand the distribution pattern of interactive and international metadiscourse features in the spoken corpus, a detail analysis was carried out. The statistical data has been shown in the table below.

Table 17: The distribution pattern of interactive metadiscourse features in the spoken corpus.

			Transition	Frame Markers	Endophoric Markers	Evidentials	Code Glosses
Text	Spoken Corpus	Count	16035	483	27598	993	5262
		% of Total	14.1%	0.4%	24.3%	0.9%	4.6%
Total		Count	16035	483	27598	993	5262
		% of Total	14.1%	0.4%	24.3%	0.9%	4.6%

As the data in table 17 shows, in the spoken corpus, endophoric markers, amounting to 24.3% of the total metadiscourse features, were the most prevalent interactive markers in the spoken corpus. The next most frequent were transitions with 14.1% of the total number of markers in the whole corpus. Code glosses with 5262 counts constituted 4.6% of the total spoken corpus, which constituted the third most frequent interactive category. With only 993 counts (0.9%), the evidentials were at the fourth position in the interactive group. Frame markers were identified as the least frequent interactive feature with 483 counts (0.4%).

Table 18: The distribution pattern of interactional metadiscourse features in the spoken corpus.

			Hedges	Boosters	Attitude Markers	Self-mentions	Engagement Markers
Text	Spoken Corpus	Count	3079	7271	1987	50110	565
		% of Total	2.7%	6.4%	1.8%	44.2%	0.5%
	5Total	Count	3079	7271	1987	50110	565
		% of Total	2.7%	6.4%	1.8%	44.2%	0.5%

Table 18 provides information on the distribution of interactional metadiscourse features in the spoken corpus. As can be seen, the self-mentions, with 5010 counts (44.2%), were the most frequently applied group of interactional metadiscourse features. They were followed by boosters (6.4%) as the second most frequent on the list in the spoken corpus. The third place belonged to hedges (2.7%) and the fourth place was taken by attitude markers (1.8%). The least frequent category was that of engagement markers (0.5%).
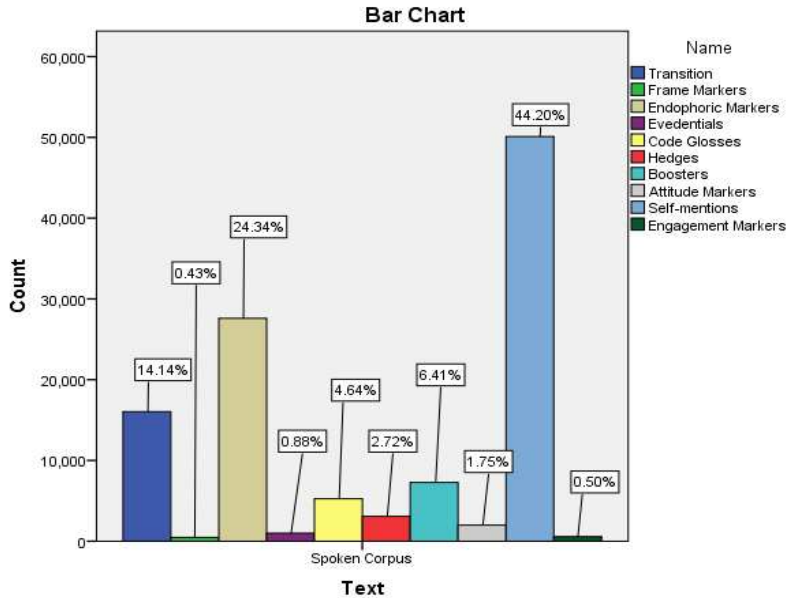
Figure 3. The bar chart of the distribution of interactive and interactional metadiscourse features in the spoken corpus.

Table 19. Cross-tabulation of the interactive and interactional metadiscourse features in the written corpus

			Type		
			Interactive	Interactional	Total
Text	Spoken Corpus	Count	50371	63012	113383
		% of Total	44.4%	55.6%	100.0%
Total		Count	50371	63012	113383
		% of Total	44.4%	55.6%	100.0%

Table 19 demonstrates the overall orientation of the interactive and interactional elements in the spoken corpus. As can be seen, the total orientation of the spoken corpus is towards interactional metadiscourse features. In other words, while interactive metadiscourse features constitute 44.40% (50371) of the

spoken corpus, the interactional group constitutes 55.6% (63012) of the total tokens identified in the spoken corpus.
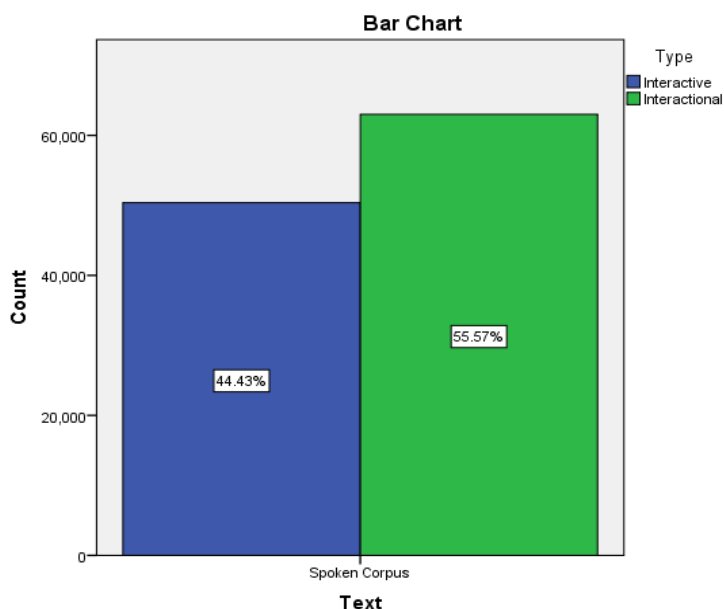
Bar Chart

Figure 4. The distribution of interactive and international metadiscourse features in the spoken corpus.

Discussion of the research problem
Response to the First Research Question
The first posed research task was to analyze the distributional pattern of interactive metadiscourse features in the written corpus. For this purpose, the frequency of interactive metadiscourse features was calculated. As the data in Figure 1 show, transitions were the most frequent interactive metadiscourse features in the spoken corpus; followed by endophoric markers and code glosses as the second and third most frequent metadiscourse features in the written corpus respectively. In the fourth and fifth position were frame markers and evidentials.

Response to the second research Questions
The second research question concerned the distribution of interactional metadiscourse features in the written corpus. As the data in table 15 can show, from among the interactional metadiscourse features, hedges were considered as the most frequent interactional items, followed by self-mentions and attitude markers as the second and third most frequent interactional metadiscourse features in the written corpus, respectively. The boosters were considered as the fourth and the the least frequent were engagement markers in the written corpus.

Response to the Third Research Question
The third research question of the study was to analyze the distribution of interactive metadiscourse features in the spoken corpus; thus, the frequency of these features was calculated. As the data in table 17 can illustrate, within the group of interactive metadiscourse features in the written corpus, endophoric markers were considered as the most frequently used items, followed by transitions as the second most frequent in the corpus. Code glosses and evidentials were the third and fourth most frequently applied. The least frequently used interactive metadiscourse group was the category of frame markers.

Response to the Fourth Research Question
The fourth research question concerned the way in which interactional metadiscourse features were distributed in the spoken corpus. On the basis of analysis it was established (cf. table 18) that self-mentions were the most frequently applied, followed by boosters and hedges as the second and third in the spoken corpus. Attitude markers and engagement markers were the fourth and fifth; interactional metadiscourse markers were found to be the least frequently used group in the spoken corpus.

5. Conclusions

Metadiscourse is a concept that is discussed and analysed frequently both in relation to written and spoken discourse (Ädel, 2010). It is regarded as a feature based "on a view of communication as social engagement and in academic contexts reveals the ways writers project themselves into their discourse to signal their understandings of their material and their audience" (Hyland, 2010, p.125). The importance of metadiscourse and its relation to the audience lies in the fact that a text or an oral presentation communicates effectively only when the writer has correctly assessed the readers' resources for interpreting (Hyland, 1999).

Analyzing selected features by applying corpus-based methodologies can unearth the distributional patterns of such features in actual contexts. Following this line of thinking, the current study focused on analyzing these features in two commercially available corpora (the British Academic Written English Corpus and the British Academic Spoken English Corpus) with the aid of the Sketch engine software.

Written Corpus
In Hyland's taxonomy, metadiscourse features are divided into two main categories: interactive and interactional. The study has shown that as far as the written corpus was concerned, the general orientation of the texts was towards the interactive category. The overall interactive orientation of the written corpus can indicate that the authors were concerned with organizing their discourse so that to predict the readers' knowledge of an explicit text as these goals are attained by

interactive metadiscourse features (Hyland, 2010). In other words, the measurement of the reliance that the authors place on interactive metadiscourse features can unearth the fact that authors are interested in consciously guiding the flow of information in such a way that they can establish their intended meanings (Hyland, 2010).

Another function of interactive metadiscourse is to help writers to make a coherent text. The prevalence of interactive metadiscourse features over interactional ones can accentuate the important function of these features as tools used for the purpose of managing the reading process. In this regard, this predominance can lend support to the claim that the texts were, to a great extent, coherent and convincing as coherence is a prerequisite for academic writing (Hyland, 2005).

In the written corpus, transitions were the most frequent category. These features are used to show contrast, concession, consequences in the course of the communication. It can be concluded that transitions are an integral part of academic writing as they can manifest how the links between various lines of argumentation work (Hyland, 2010). Their distribution also reflects the fact that the authors made spaces for alternative ideas, reasoning, and claims in their discourse and these are among characteristics of academic writing. Indeed, by using transitions, writers help their readers to interpret the links that exist between relevant ideas, and, in another perspective, help to unearth the reasoning of the authors in an unambiguous way.

However, the second most frequent group of interactive metadiscourse features in the written corpus were endophoric markers. These features are used to show how writers cite and refer to other parts of their texts. By using these elements, the reader can recover the intended meaning of the author and gain a better comprehension of the total message (Hyland, 2005). As a matter of fact, referring to other parts of the text is a unique feature of academic discourse and authors try to refer to illustrations, examples, sections, parts, and arguments in other parts of their texts, which is a feature typical of academic discourse (Hyland, 2002).

The third most frequent group of interactive metadiscourse features were code glosses that are used to provide the readers with additional information and make sure that the reader has unfolded the meaning intended by the author(s). These are part of academic writing in that readers can reflect writer's predictions towards the level of understandability of the text on the behalf of the readers (Hyland, 2002). The usage of glosses can manifest that the authors might not have been sure that their intended proposition(s) had been fully received and understood by the receivers; making them use glosses to facilitate the process of learning .

Evidentials were the fourth most frequently used category in the written corpus. They are used to refer to other resources out of the writer's message. Citations are common in academic writing, and academic texts, which can show that the authors have carefully read other sources. The fact that these features had a much lower distribution in the corpus, lower than other metadiscourse features, may be an indication that the authors were convinced of their propositions and did not have

to cite others to justify their argumentation producing more coherent texts (transitions).

In the written corpus, the least frequently used interactive metadiscourse features were frame markers. The fact that they were rarely used may indicate that writers of academic texts were not signalling text boundaries or sequencing different parts of the discourse as they are attained by these features (Hyland, 2005).

As far as the interactional metadiscourse features in the written mode were concerned, the most preponderant feature was hedging. In academic writing, it is very important to distinguish facts from unsupported ideas and hallucinations. This can be assured by the use of hedges. The extra usage of hedges can support the fact that writers try to assess their intentions in such a way that they are acceptable and academically persuasive to the writers (Hyland, 2005). These features can manifest that the statements or the propositions offered by authors in the text are mostly based on their own interpretation rather than some certain amount of knowledge.

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

The third most frequently applied group of interactional metadiscourse features in the written corpus were attitude markers. These features are used by authors to show their attitudes towards an argument, a proposition, and a claim. In other words, they are used to show effective not the scientific and logical attitude of the authors (Hyland, 2005). In academic discourse, there is less space for using such effective elements as these features are mostly used not in academic discourses, but in literature and poetry. As a result, it can be claimed that the authors of the written corpus did not have interest in stepping out of the scientific boundaries.

The fourth place belonged to boosters. These features are used to show a degree of certainty and to close down an argument, making no space for alternative propositions (Hyland, 2010). With a low frequency of boosters the hedges (uncertainty elements) can reveal the fact that authors lack confidence in most of their ideas, propositions, claims, and arguments; eschewing them to express their 100% certainty in the context.

The least frequent group of features in the academic written corpus were engagement markers. Usually, engagement markers are used to explicitly attract attention of the reader towards a proposition (Hyland, 1998). On the other hand, scientific texts are not the kind of register in which direct instructions for the readers can be found. By using such tokens in a rare mode, a person may suggest that the authors underestimate the presence of the readers during the interaction.

Spoken Corpus

The quantitative analysis showed that, similarly to the written corpus, the overall tendency was that speakers were interactive-function oriented. As far as the interactive category of metadiscourse features were concerned, in the spoken corpus, unlike the written corpus, endophoric markers were the most frequent group, not the transitions. As said above, endophoric markers are used to refer to other parts of the texts. This can add extra support to the claim that although making a speech coherent and well organized is an important aspect of spoken language, and speeches (transitions), more emphasis is put on referring to additional materials to facilitate the understanding of the message on the behalf of the reader. This is usually done because, in the spoken corpus, there are no written materials, unlike the written context, through which the reader can easily flashback to them; therefore, the author himself should carry this to help the reader not to get distracted from the speech, or get off the right track. Like in the written corpus, the third, fourth and fifth most frequently used interactive metadiscourse features were code glosses, evidential and frame makers, respectively.

In the interactional category of metadiscourse features of the spoken corpus, self-mentions were the most preponderant among all the interactive metadiscourse features. This means that unlike the written corpus in which hedges were the most frequently applied interactive metadiscourse features, in the spoken corpus authors used self–mentions more often, showing their strong presence in the text. The usage of self-mentions is an indication of the fact that the authors had self-representation and projected not only themselves, but also their claims about propositions (Ivanic, 1998).

Boosters came second, unlike the written corpus in which boosters were in the fourth position. This means that the authors clearly project their ideas and claims. The combination of both self-mentions and boosters, can reveal that the authors in the spoken corpus not only tried to "project" themselves in the course of a lecture or a presentation, but also tried to clearly express their claims, ideas, findings, and beliefs. This extensive usage of self-mentions and boosters prevented the authors from using hedges in the spoken corpus. While the boosters show the certainty of the authors, the usage of hedges represents uncertainty. This can show that while the authors in the written corpus were less certain and more dubious about claims or proposition, in the spoken corpus, it was exactly vice versa.

Attitude markers and engagement markers were the least frequently used interactional metadiscourse features in the spoken corpus. This can show that the authors were not willing to express their affective attitude towards the propositions as this is done through the usage of attitude markers. The engagement markers were not used extensively, which means that authors underestimated the presence of their audience in the discourse space. This is in sharp contrast to self-mentions in which authors show their strong presence; and inevitably, they have to downplay the presence of the audience.

Overall, both corpora were interactive-tokens oriented, which means that the authors were preoccupied in organizing their product (either spoken or written) in such a way that the target audience could follow a well-established and coherent course of the message. In the written corpus, the sequence of interactive metadiscourse features were transitions, endophoric markers, code glosses, evidentials, and frame markers, respectively. In the spoken corpus, however, the sequence of metadiscourse features was endophoric markers, transitions, code glosses, evidentials and frame markers.

In the interactional category of metadiscourse features, the written corpus had the sequence of hedges, self-mentions, attitude markers, boosters and engagement markers, respectively. However, in the spoken corpus, the sequence was based on self-mentions, boosters, hedges, attitude markers and engagement markers.

Implications, Limitations, and Suggestions for Further Studies

This research has some implications for future origrammes. One implication is for researchers in the field of Corpus Linguistics. Those interested in this area of inquiry can use this study in order to be able to perform similar programmes. Another implication is for those who are interested in comparative studies, which formed a significant part of the present one. In addition, those interested in metadiscourse features can benefit from this study.

Clearly, this research had some limitations. One limitation was that the number of data in the corpora was different; it would have been better if the texts had been similar in terms of volume. Moreover, some metadiscourse futures may have been neglected in the course of data extraction. In addition, the researcher had no control over the corpora design and data gathering as the corpora were both commercially available. The last limitation of this research was the fact that this research was focused on and limited to the academic genre and did not take other genres into consideration.

Hopefully, this research may spark off new studies. This study was limited to Hyland's model of metadiscourse features, which provided its theoretical framework. It is suggested that metadiscourse features should be analyzed with the use of other categories of metadiscourse features contributed by other theoretical models. Moreover, the present study was focused on analyzing metadiscourse features solely in academic context (genre). It can be a good idea that other genres, such as fiction, news, law, etc. should be analyzed in terms of metadiscourse features. The current study was a synchronic research, a diachronic perspective could undoubtedly be promising and allowed to track the metadiscourse features and their development in different time spans.

References

Abdollehzadeh, Esmail. 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.

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

Ädel, Annelie. 2006. *Metadiscourse in L1 and L2 English*. Amsterdam: John Benjamins. http://dx.doi.org/10.1075/scl.24.

Ädel, Annelie. 2010. "Just to give you kind of a map of where we are going A taxonomy of metadiscourse in spoken and written academic English". *Nordic Journal of English Studies*, Vol.9, No. 2, pp. 69-97. https://doi.org/10.35360/njes.218

Ahour, Touran., & Entezari Maleki, S. 2014. „The effect of metadiscourse instruction on Iranian EFL learners' speaking ability". *English Language Teaching*, Vol. 7, No. 10, pp. 69-75. https://doi.org/10.5539/elt.v7n10p69

Amiryousefi, Mohammad., & Eslami Rasekh, A. 2010. "Metadiscourse: Definitions, issues and its implications for English teachers". *English Language Teaching*, Vol. 3, No. 4, pp. 159-167. http://dx.doi.org/10.5539/elt.v3n4p159

Bailey, Stephen. 2003. *Academic writing: A handbook for international students* (1st ed.). Abingdon: Routledge.

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. TEFL.* Unpublished Master's Thesis, University of Tehran, Tehran.

Boggel, Sandra. 2009. *Metadiscourse In Middle English And Early Modern English Religious Texts: A Corpus-Based Study*. Frankfurt am Main: Peter Lang.

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

Brown, G., & Yule, G. 1993. *Discourse Analysis* (1st ed.). Cambridge: Cambridge University Press.

Camiciottoli, Belinda Crawford. 2003. "Metadiscourse and ESP reading comprehension: An exploratory study. *Reading in a Foreign Language*, Vol. 15, No. 1, pp. 1-16. Retrieved from http://nflrc.hawaii.edu/rfl/April2003/camiciottoli/ camiciottoli.html (accessed 2019)

Chafe, Wallace. 1982. "Integration and involvement in speaking, writing and oral literature". In: D. Tannen (ed.), *Spoken and written language: Exploring orality and literacy* Norwood, NJ: ABLEX, pp. 35-53.

Cheng, X., & Steffensen, M. S. 1996. "Metadiscourse: A technique for improving student writing. *Research in the Teaching of English*, Vol. 30, No. 2, pp. 194-181. https://doi.org/10.1177/0741088393010001002

Cissna, Kenneth. N., & Anderson, R. C. 2002. *Moments Of Meeting: Buber, Rogers And The Potential For Public Dialogue*. Albany, NY: SUNY Press.

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

Crismore, A., Markkanen, R., & Steffensen, M.S. 1993."Metadiscourse In Persuasive Writing: Study of Texts Written by American and Finnish University Students". *Written Communication, 10* (1), pp. 39-71.

Crismore, Avon., & Abdollahzadeh, E. 2010. "A review of recent metadiscourse studies: The Iranian context". *Nordic Journal of English Studies*, Vol. 9, No. 2, pp. 195-219. https://doi.org/10.35360/njes.223

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.

Flowerdew, John & Steve Tauroza. 1995. The effect of discourse markers on second language lecture comprehension. *Studies in Second Language Acquisition* 17, pp. 435-458. https://doi.org/10.1017/S0272263100014406

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. Theory and Practice in Language Studies, Vol. 5, No. 2, pp. 309-317. doi:10.17507/tpls.0502.10.

Ghahremani Mina, K., & Biria, R. 2017. "Exploring interactive and interactional metadiscourse markers in discussion sections of social and medical science articles". *International Journal of Research in English Education*, Vol. 2, No. 4, pp. 11-29. doi:10.29252/ijree.2.4.11.

Harris, Zellig. 1959. "The transformational model of language structure". *Anthropological Linguistics*, Vol.1, No. 1, pp. 27-29.

Hudson, R. A. 1980. *Sociolinguistics* (1st Ed.). Cambridge: Cambridge University press.

Hyland, K. 1998. "Persuasion and context: The pragmatics of academic metadiscourse". *Journal of Pragmatics*, Vol. 30, pp. 437-455. http://dx.doi.org/10.1016/S0378-2166(98)00009-5.

Hyland, Ken. 1999. "Talking to students: Metadiscourse in introductory course books". *English for Specific Purposes*, Vol. 18, No. 1, pp. 3-26. https://doi.org/10.1016/S0889-4906(97)00025-2

Hyland, Ken. 2002. "Authority and invisibility: authorial identity in academic writing". *Journal of Pragmatics* Vol. 34, pp. 1091–1112. https://doi.org/10.1016/S0378-2166(02)00035-8

Hyland, Ken. 2004. "Disciplinary interactions: metadiscourse in L2 postgraduate writing". *Journal of Second Language Writing*, Vol. 13, No. 2, pp. 133-151. doi:10.1016/j.jslw.2004.02.001

Hyland, Ken. 2004. "Patterns of engagement: Dialogic features and L2 student's writing". In: L. Ravelli & R. Ellis (eds.), *Academic Writing in Context: Social-functional Perspectives on Theory and Practice*. London: Continuum.

Hyland, Ken., & P. Tse. 2004. "Metadiscourse in academic writing: A reappraisal". *Applied Linguistics*, Vol. 25, No. 2, pp. 156-177. https://doi.org/10.1093/applin/25.2.156

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

Hyland, K. 2010. "Metadiscourse: 'Mapping interactions in academic writing'". *Nordic Journal of English Studies* 9: 2, pp. 125-143. https://doi.org/10.1093/applin/25.2.156

Latawiec, Bogusław. 2012. *Metadiscourse in oral discussions and persuasive essays of children exposed to collaborative reasoning* (Unpublished doctoral dissertation). University of Illinois, Urbana-Champaign.

Ivanic, Rosalind, 1998. *Writing an Identity: The Discoursal Construction of Identity in Academic Writing*. Amsterdam: Benjamins. https://doi.org/10.1075/swll.5

Jalilifar, Alireza. & M. Alipour. 2007. "How explicit instruction makes a difference: Metadiscourse markers and EFL learners' reading comprehension skill". *Journal of College Reading and Learning* 38: 1, pp. 127-148.
https://doi.org/10.1080/10790195.2007.10850203

Kong, R., & Xin, X. 2009. "Empirical study on metadiscourse in Chinese EFL learners' oral communication". *CELEA Journal* 32: 1, pp. 52-64.

Kopple, William. J. 1985. "Some exploratory discourse on metadiscourse". *College Composition and Communication* 36: 1, pp. 82-93. doi: 10.2307/357609.

Kruger, A. 2004. "Corpus-based translation research: its development and implications for general, literary and Bible translation". *Acta Theologica* 22: 1, pp. 70-106. doi:10.4314/actat.v22i1.5455.

Love, R., C. Dembry, C., A. Hardie, V. Brezina & T. McEnery. 2017. "The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations". *International Journal of Corpus Linguistics*, Vol. 22, No. 3, pp. 319-344. doi: 10.1075/ijcl.22.3.02lov

McEnery, Tony, & Andrew Hardie. 2012. *Corpus linguistics: Method, theory, and practice* (1st ed.). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511981395

McGillivray, Barbara & A. Kilgarriff. 2013. "Tools for historical corpus research and a corpus of Latin". In: Bennett, P., Durrell, M., Scheible, S., & Whitt, R. J. *New methods in historical corpora*. Tübingen: Narr.

Mukherjee, Jogeshwar. 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* . Wien: Peter Lang, pp. 5-24.

Penz, Hermine., E. Maria Graf, & G. Marko. 2016. Verbal workshop: metadiscourse in spoken language". Retrieved from https://static.uni-graz.at/fileadmin/gewiinstitute/Sprachwissenschaft/ OELT_2016/VERBAL_Workshop_Metadiscourse_OELT_2016.pf (accessed 2019)

Pérez, Marta. A., & Elisabet A. Macià. 2002. "Metadiscourse in lecture comprehension: Does it really help foreign language learners?" *Atlantis Journal* 24: 1, pp. 3-21.

Schiffrin, Deborah. 1980. "Metatalk: Organisational and evaluative brackets in discourse. Sociological inquiry". *Language and Social Interaction* 50, pp. 199-236. https://doi.org/10.1111/j.1475-682X.1980.tb00021.x

Thompson, S. E. 2003. Text-structuring metadiscourse, intonation and the signaling of organization in academic lectures. *Journal of English for Academic Purposes* 2, pp. 5-20. https://doi.org/10.1016/S1475-1585(02)00036-X

Tymoczko, Maria. 1998. Computerized corpora and the future of translation studies. *Meta: Journal des traducteurs* 43: 4, pp. 652-660. doi: 10.7202/004515ar.

Vahid Dastjerdi H., & M. Shirzad. (2010). The impact of explicit instruction of metadiscourse markers on EFL learners' writing performance. *The Journal of Teaching Language Skills* 2: 2, pp. 155-174.

Weisser, M. 2016. *Practical Corpus Linguistics: An Introduction To Corpus-Based Language Analysis*. Oxford: Wiley Blackwell. https://doi.org/10.1002/9781119180180

Zanettin, Federico. 2012. *Translation-Driven Corpora: Corpus Resources For Descriptive And Applied Translation Studies*. Place of publication not identified, UK: St. Jerome Publishing.

Zarei, Gholam & Sara Mansoori. 2007. Metadiscourse in academic prose: A contrastive analysis of English and Persian research articles. *The Asian ESP Journal* 3: 2, pp. 24-40.