An Investigation of Evaluative Prosody in the Collocations Used in Introduction Sections of Scientific English Articles of Hard and Soft Disciplines

Leila Ghaemi, Ph.D. Candidate, Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran

Leili_gha2004@yahoo.com

Nesa Nabifar, Assistant Professor, Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran

nesanabifar13@gmail.com

Davud Kuhi, Assistant Professor, English Language Department, Islamic Azad University, Maragheh Branch, Maragheh, Iran

davudkuhi@yahoo.com

December 12, 2022 February 28, 2023

ABSTRACT

This study investigates the collocations, as specific type of multiword sequences that frequently occur in the scientific writing of hard and soft disciplines based on Hoey's (2005) model. The study also extends the semantic prosody of collocations in general words across disciplines. To this end, the introduction sections of 1000 research articles (RAs) from two disciplines of soft and hard were gathered from disciplines of applied linguistics, sociology, and psychology under the category of soft science and from the majors of computer science, chemistry, physics and medicine under the category of hard science in order to find a relationship between these two types of RAs with differences in discipline and the authors' use of collocations and their semantic prosody. In this study, function words containing articles were ignored that their rates and frequencies were higher than the content words, including specific and general words. The findings indicated that the collocates of general words expressed the positive, negative, and neutral senses and being in soft or hard science did not affect their semantic prosody. It means that the general words directly expressed the meaning they transferred to the reader, without the interference of the context. The results can result in collocation teaching, which is a popular genre of academic writing assigned across disciplines as well as in language-related courses.

***Keywords*:** collocations, evaluative prosody, hard and soft disciplines, introduction sections

Introduction

In formal essays, academic writing is a specific style that allows authors to use formal terminology and to touch on the conceptual limits of their disciplines in the fields of their special knowledge. When it

comes to writing, especially academic writing, new members also make the use of vocabulary suitable for difficult academic documents (Tribble, 2011). To overcome this problem, instructors are interested in new instructional instruments to inform the implementation of academic writing skills for their students. The use of corpus analysis is one such technique that focuses on different disciplines or even with comparing non-native and native authors in the big field of academic or scientific writing. The value of formulaic multiword sequences, like collocations in a number of academic registers, has progressively been documented by corpus-based research (Xiao, & McEnery, 2006) as an essential feature that aims to develop a rational and ordered stream in the written expression and also a coherent and comprehensive whole, it generates cohesiveness. The use of formulaic sequences (FS) in total and collocations in specific can play a significant role in academic discourse in this regard. Collocation is a linguistic phrase coined by Firth (1975), who asserts that you can express a term by the association it keeps. The basis of lexical items being recognized as collocations, according to Hoey (1991), is that they appear together with greater than random probability in their (textual) context.

Davis and Morley (2015) say that by helping learners to structure their insights and enhance their composition style, the identification of series such as collocations can aid. Language is repetitive in nature to a great degree (Meunier, 2012) and understanding and developing FS, frequently applied in academic literature, can be a strong baseline for second language students to coherently express their messages and ideas. In addition, learning collocations allows research paper writers mindful of the expressiveness of language. Based on Hoey's Lexical Priming Theory (Hoey, 2003, 2004, 2005), lexis and grammar functions are reversed in such a way that lexis is organized in a complicated and organized way and that grammar is a result of this lexical arrangement. This recent approach has the potential to explain how naturalness is done and how an understanding of what is natural can influence interpretations of what is feasible.

Collocation is a significant aspect in genuineness in this regard (Hoey, 2005). The origin of collocation contains an understanding for Hoey (2004), who proposes the priming theory: the only argument that appears to explain for the presence of collocation is that each lexical word is primed for collocational usage. In addition to the importance of the collocations and their position in scholarly literature, the writers and English learners may accomplish a controversial definition. Howarth (1998) believed that because of their absence of collocational skills, English as a Foreign Language (EFL) and English as a Second Language (ESL) students encounter specific difficulties in generating suitable idiomatic phrases. Gilquin and Paquot (2008) also argued that L2 students are struggling with the adequate use of FS, rendering their academic writing inadequate and too colloquial. These results indicate that ESL/EFL students ought to be subjected to substantial and commonly utilized FS or sentences used in academic writing to improve collocational awareness in overall.

Semanticc prosody, also known as evaluative prosody, is a concept that applies to collocations that has gained popularity in corpus-based investigations. Semntic prosody is characterized as a type of meaning generated by the vicinity of a set of collocates that expresses the author's or speaker's mood, stance, or position (Louw, 2000). As per the process of phonological coloring, Louw (1993) coupled semantic prosody to phonological prosody postulated by Firth (1957). The semantic prosody of a form is characterized as a regular aura of meaning infused by its collocates. Evaluative prosodies were depicted as representations of either pejorative or ameliorative semantic changes over time, in line with the extremely common types that may diverge into good' and bad. The diachronic aspect of semantic prosodies was also addressed by Louw (1993), who stated evaluative prosodies are the culmination of a span of refining via historical development.

As an essential aspect of collocation, semantic prosody can reflect the narrator or writer's perspective, posture, standpoint, or thoughts regarding the things and concepts being discussed (Hunston & Thomsom, 2000). Parallel to Louw's (1993) aura of meaning, Bublitz (1996) stated phrases can have a picture or a halo that can be favorable, pleasant, nice, unpleasant, negative, or awful, and that evaluative prosody relates to positively or negatively semantic shading of collocate and node. Furthermore, semantic prosody is identified as critical element in speech selection and plays a decisive part in the lexical grammar co-selection process (Stubbs, 2009). As a result, evaluative prosody offers a novel region for lexicography and lexical semantics, as well as a macro lens for observing, studying, and explaining lexical function. Semantic prosody has been described as the manner in which words in a manuscript might collocate with a relevant collection of words, sometimes indicating unseen mindsets (Baker, Hardie, & McEnery, 2006). In a handful of sociocultural contexts of involvement, such as text, lexicography, pragmatics, interpretation, and discourse analysis, as Dam-Jensen and Zethsen (2008) claimed, study results in semantic prosody will pursue to supply more research in broad areas of language-ralted studies in a multitude of sociocultural contexts of involvement.

These twenty years have seen a large systematic review dedicated to semantic prosody (e.g. Bi, 2019; Hu, 2015; Hunston, 2007; Selmistraitis, 2020; Du, Wang, & Yang, 2018; Wei & Li, 2014; Xiao & McEnery, 2006) that most of them learned close synonyms in various corporations, encompassing the Corpus of Contemporary American English (COCA) scholarly texts and with native and non-native people. To date, any of the researchers studying semantic prosody have not chosen the target items like collocations from a specific academic list obtained from two disciplines of hard and soft corpus data. Researchers around the world are already under enormous pressure within academia to read and write vast quantities of scholarly texts. English is undeniably one of the key mediums of academic texts, although academic clients are certainly not native English learners, in addition, collocations are under-researched and overlooked and the way in which Prime collocates with each other across the disciplines. To ignore studies in terms of semantic prosody with collocations across disciplines is one barrier that may obstruct non-native students on the road to academic gain, so the present study aimed to study the evaluative prosody with the high frequently used collocations across various academic sciences, with an effort to empower the skillful outcome of foreign academic users by distinguishing the patterns of those collocations with special focus on semantic prosody. In addition, expanding the understanding of semantic prosody and collocation by students can be advantageous for vocabulary teaching (Bi, 2019). As a new area in lexicology research, collocation and semantic prosody may resolve the problem of the proper use of collocations or formulaic sequences with corporate aid throughout disciplines or sciences. Without doubt, semantic prosody and collocations could differ in the introduction sections of scientific research articles written by authors of different disciplines. Hence, the current research, as a corpus-based study, used a contrastive analysis of collocation and semantic prosody based on Hoey's lexical priming framework (2005) in the introduction sections of research articles (RAs) written by soft and hard disciplines. According to the research aims, the following research questions are addressed in this study:

1. How frequently are collocations employed in the introduction sections of hard and soft science research articles based on Hoey's (2005) lexical priming model?

2. How do hard and soft science authors of research articles use shared collocations (those found in both corpora) in terms of evaluative prosody for co-occurring semantic items?

This section elaborates on two key-terms of collocations and semantic prosody tht frequently were used in this study.

Collocation

Collocation has been researched for about a half-century. Firth (1957) planned to bring forward implement the criteria of collocability (Firth, 1957). Collocations of a specific term, thus according Firth (1968), are declarations of the word's usual or typical positions. Firth's collocation idea is essentially quantitative (Krishnamurthy, 2000). Many corpus linguists advocate for the statistical approach of collocation, namely, for instance, Hoey (1991), Sinclair (1991), Halliday (1966), Greenbaum (1974), Stubbs (1995), McEnery and Wilson (2001), and Hunston (2007). Collocations of a specific word, thus according Firth (1968), are declarations of the word's usual or typical positions. Firth's collocation idea is essentially quantitative (Krishnamurthy, 2000). Many corpus linguists advocate for the statistical approach of collocation. Whilst Greenbaum's notion does not specify how regular it should be to recognize the happening of two lexical elements as a collocation, Hoey (1991) only utilized the term collocation each time a lexical item represents in its (textual) form with other things with a higher than randomly selected.

Co-occurring of items directly related to the notion of lexical priming, a newly developed Hoey's (in particular Hoey, 2003, 2004a, 2005) linguistic theory that attempts to apply corpus-linguistic constructs like colligation and collocation to the analytical outcomes of psycholinguists engaging in word association acceleration and retardation. Since the mid-1950s, linguistic theories have differed in the degree to which the syntactic reasons for uncertainty were relevant rather than in the essence of disambiguation (Hoey, 2005). Most notably, they have generally lacked the ability from a corpus-linguistic perspective to accommodate for collocation (that can be narrowly described as the partially random propensity of words to typically occur in the setting of each other).

Semantic Prosody

According to Stubbs (2002), there are often semantic links across collocates and node, mostly amongst the collocates, when moving from form to sense. Semantic prosody, a sort of feeling produced by the likeness of a defined set of collocates, can be referred to as the collocational importance coming from the interplay between a given node and its normal collocates (Louw, 2000). There can be semantic prosodies in both single words and phrases (Schmitt & Carter, 2004). The main role of semantic prosody is to convey the mood or assessment of the communicator (Louw, 2000). Usually, evaluative prosodies are not positive, with some of them having a capably beneficial value. A communicator, nevertheless, may also breach an evaluative prosody requirement to obtain some influence on the audience, such as insincerity, irony, or satire can be clarified by recognizing evaluative prosody deviations (Louw, 1993).

It is a subject of debate whether evaluative prosody is a kind of conventional meaning. Partington (1998) and Hunston (2002) seem to assume that evaluative prosody is connotational, whilst Louw (2000) asserts that evaluative prosodies are not solely connotational, because the power driving evaluative prosodies is more firmly collocational than the schematic facets of connotation. In Stubbs' (2002) statement stated above, the notion coming from the shared embedding of the collocates of a given node word can be attributed to semantic preference that is described as semantic preference by a lexical group of often appearing collocates sharing semantic attribute (Stubbs, 2002). For instance, Stubbs (2001) claims that huge collocates with items from the same semantic collection, implying 'amounts and quantities' (e.g. size, number(s), part, number, quantity(s)), whereas Partington (2004) claims that change and absence of state is a prevalent aspect of maximizer collocates like definitely, wholly, fully, and totally.

Many scholars have investigated the effect of evaluative prosody throughout the last twenty years, namely Louw (1993), Stubbs (2002), Partington (2004), Schmitt and Carter, 2004, Sinclair (2004), Hunston (2011), and many more. Along, these experiments have shown that the attitudinal value of an object is recognized Unless it is utilized and this definition is often featured as neutral, positive or negative in the sense of its normal collocations (e.g., Xiao & McEnery, 2006). In order to demonstrate, Louw (1993) named numerous things, including absolutely, symptomatic of, bent on, and victim of, that appear to be correlated with unfavorable meanings. Conklin and Schmitt reported on the other aspects that align with positive connotations, such as the provision of information and services (2008). These years have experienced a great number of researches dedicated to evaluative prosody, as represented by the efforts of researchers such as Sinclair (1991), Sardinha (2000), Partington (2004), and Xiao and McEnery (2006), but contrastive analysis between two sciences of hard and soft in the introduction sections of scientific research articles studying both frequency-driven approach to collocations and contrastive analysis of semantic prosody is still under-explored. Moreover, unlike previous studies which adopted whole RA approach, this study delves into the use of collocations and evaluative prosody in introduction sections of hard and soft sciences because writing introductions, based on Swales and Feak (1994), is widely thought to be difficult and troublesome, and writing a solid introductory section typically appears to be a struggle won hard.

The reality that the presentation is supposed to get a justification for the realization of the project study by specifying the aims and importance and to attract readers by generating interest in the subject may be what causes writing Research Article Introductions (RAIs) relatively more difficult (Swales & Feak, 1994). Therefore, one of the important factors deciding whether or not the RA is likely to be published is possibly the efficacy of writing the Introduction portion. The current investigation thus aimed at contrasting collocations used in the Introduction parts of hard and soft disciplines in order to gain understanding on the difference between academic disciplines in accordance to their traditional and appropriate collocational features and their semantic prosody. The choice of the two hard and soft science corporations was based on a common typology for the categorization of academic knowledge that distinguishes between 'pure' or 'applied' and 'hard' or 'soft' domains. Consequently, we internet to research on the basis of this category on texts of medicine, physics, and so on that are the optimal indicative of the hard academic discourse, and on texts of psychology and applied linguistics that serve as the embodiment of soft academic discourse. In view of this, this paper conducted a contrastive analysis of collocations in accordance to the frequency and semantic prosody between the authors of scientific research papers in hard and soft sciences, using a corpus-based approach.

Methodology

Corpus

The data for the current research was obtained from the introduction section of 1000 RAs of two disciplines of soft and hard. The corpus of this study was gathered from disciplines of applied linguistics, and psychology (soft science) and computer science, physics and medical science (hard science) in order to find a relationship between these two types of RAs with differences in discipline and the authors' use of collocations and their semantic prosody. A significant genre is considered in any research article in the academic community. As Hyland (2000) has noted, RA simultaneously pursues binary purposes: the first objective deals with the propagation of new knowledge to the adherents of their discourse culture, while

the second objective is to convince the members of the discourse to embrace the claims. To ensure the generalizability of the results to the target discourse and account for representative practices of discourse community in different disciplines, leading journals in both hard and soft sciences were selected based on consultation with discipline experts and the Impact Factors (MIFs) ranging from 1 to 5 that were reported in the Journal Citation Reports (JCR) in 2015.Five data bases including Elsevier, Sage, Springer, Taylor & Francis, and Wiley Online Library comprise the sources of RAs in two disciplines published between 2010 and 2020. Research articles were randomly selected from each database, yielding a corpus of 1000 research articles of approximately 603,000 words (soft science comprising of 288,582 words) and (hard science comprising of 314,471 words). In the selection of this corpus, we ensured that there is a proportionate number of native and non-native writers. Tables 1 and 2 shows the corpus description of research articles in soft science and hard science along with the Impact Factors and the number of introduction sections.

Tables 1

Corpus Description of Research Articles in Soft Science

Journal Title		Impact Factor	Number
Psychology	Journal of Memory and Language	3.893	122
	New Ideas in Psychology	1.550	48
	Journal of School Psychology	2.981	50
	Acta Psychologica	1.380	30
Applied Linguistics	Teaching and Teacher Education	2.686	50
	Language Learning	3.408	30
	The Modern Language Journal	3.077	43
	Journal of Second Language Writing	3.538	55
	Learning and Instruction	3.323	72

As Table 1 shows, two disciplines of Psychology and Applied Linguistics were randomly chosen as majors in soft science. Different journals such as Journal of Memory and Language, New Ideas in Psychology, Journal of School Psychology, and Acta Psychologica with IFs ranged from 1.3 to 3.8 were the focus of Psychology with the total of 250 articles. For Applied Linguistics, the journals of Teaching and Teacher Education, Language Learning, The Modern Language Journal, Journal of Second Language Writing, Learning and Instruction were chosen, which their numbers were 220 and the IFs ranged from 2.6 to 3.5. Table 2 shows the description of corpus in hard science.

Tables 2

Corpus Description of Research Articles in Hard Science

Journal Title		Impact Factor	Number
Computer	Ad Hoc Networks	3.645	40
	Future Generation Computer Systems	6.125	40
	Expert Systems with Applications	5.452	50
Medicine	Clinical Investigation	2.686	50
	Cardiovascular Interventional Radiology	2.75	11
Other	American Journal of Alzheimer's Disease &	2.144	45
	Cancer Genetics	3.105	20
	Dementias	1.614	20
	Radiotherapy and Oncology	4.856	30
	Journal of Cerebral Blood Flow & Metabolism	5.681	30

	Journal of Computational Physics			
Physics			52.98	20
	Technometrics	6		
				40
			2.091	

Based on the descriptions of Table 2, the three majors of Computer, Medicine, and Physics were used as hard science majors. It is worth noting that the number of journals in medicine was higher than the other majors that the reason go to the fact that during data gathering, the researcher found that the number of words in two majors of Computer and Physics is very low, so she obliged to ignore most of the collected articles and focus mostly on Medicine research articles that have standard introduction sections.

Data Collection

To analyze the corpus, the introduction sections were identified in the articles and stored separately. The articles that comprised of isolated heading or sections such as introduction, literature review, method, and so were selected and those articles, which the introduction sections were embedded in the review section and vice versa were excluded from the analysis and converted into plain text files (txt). Afterwards, using AntConc software (Anthony, 2011), we identified the frequency of collocations in both disciplines. The reason for choosing the introductory portions of RAs was that introductions, as a significant element for knowledge communication, have gotten more attention in latest years of the academic world's information boom. According to Swales and Feak (1994), writing introductions is complicated and problematic, or what tends to make writing Research Article Introductions (RAIs) relatively more challenging is that the introduction is predicted to include a justification for undertaking the study by asserting the targets and importance, and to generate attention by producing enthusiasm in the topic (Swales & Feak, 1994). According to Martn, Rey-Rocha, Burgess, and Moreno (2014), the Introduction is the most rhetorically complex section after the Discussion section, and hence the most complicated to write, which may explain why journal editors and reviewers scrutinize and harshly condemn these two parts of the RAs that were presented for publishing. As a result, the quality of the Introduction section is undoubtedly one of the most crucial factors in determining whether or not the RA will be published. Furthermore, two disciplines of hard science and soft science were studied. In the context of knowledge, a discipline refers to a specific field of academic study (such as applied linguistics and medical science) and disciplinary variation, therefore, refers to differences inherent in different disciplines, in this case, in terms of lexical priming model of Hoey (2005). It is important to note that these are slang phrases for comparing scientific subjects based on perceived methodological rigor, exactitude, and objectivity. Natural sciences are "hard," according to Wilson (2012), but social sciences are typically defined as "soft."

After that the necessary data gathered in collocations, it was analyzed by lexical priming model that accounts for collocation since as Hoey (2005) stated each term is mentally linked for collocational use. In accordance with the model, for analyzing the collocations, the noun group was selected to be compared for its collocational (the word or words that characteristically accompany a term) patterns across the

corpora. After above phases, the rates and frequency of the collocations in both disciplines were presented in tables.

The other focus of this study was on analyzing the data based on semantic association such as semantic prosody. As Hoey (2005) stated Semantic relationships go forward by enabling the grouping of colloquates into semantic pairs. These show language patterns which go beyond collocation with a set of variable elements to form phraseological units. Using the example of 'consequence' that absorbs adjectives that offer a 'logical' connection, the argument for the semantic sets is formed and can be divided into sub-groups of 'necessity', directness or level of the systematic argument and the naturalness or expectation of the system. For exploring the semantic association such as semantic prosody, the researcher scrutinized on the high frequently used collocations in both hard and soft sciences and after that she manually analyzed the occurrence of that word with the other word or words making association semantically and after that she reported the data in the form of qualitative data analysis.

Data Analysis

The current research as both qualitative and quantitative data analysis, tried to gain the aim of the study. The first research question was reported in terms of frequency and to find the existence of any meanigful differences between two kinds of research articles in terms of the rates of collocations, the descriptive statistics containing frequency and percentage was used and after that the results were compared via Chi-Square data analyses. In terms of the second research question, semantic prosody in each corpus was investigated and it was compared with each other through descriptive data analysis. To ensure the reliability of the analysis in the data categorization process, 10% of the data was rechecked and reanalyzed independently for collocations and semantic prosody by a second researcher (a Ph.D. graduate of TEFL) who was briefed about the purpose of the study by the researcher. Also, the field of study of this expert was discourse analysis, and she was familiar with the data analysis phase. The second rater coded 10% of the data, taken randomly from the corpus and finally, the inter-rater reliability was estimated and reported. The inter-rater agreement, measured using Cohen's Kappa formula, was found to be Kappa = 0.929, $p = 0.000$.

Results

As already stated, based on Hoey's model (2005), for analyzing the collocations, the noun group was selected to be compared for its collocational patterns across the corpora, to this end, function words were ignored in the current study and just the highly frequent nouns, no difference in their specificity or generality, were chosen from corpus. Based on the results of data analysis, in both corpora, the function words such as “the, of, and, in, to, a, as, for” were used in high frequency. The high frequent content words in hard and soft disciplines are shown in Table 3 below.

Table 3

The Top 5 Collocations in Soft and Hard Disciplines

Soft science/ token: 9416		Hard science / token: 311550	
Collocates	Tokens	Collocates	Tokens

Language	2164	Studies	7770
Writing	1583	Treatment	6330
Learning	1336	Data	5940
Research	1310	Method	5570
Students	1214	Patient	1560

As above table shows, in soft science, the five content words of *Language*, *Writing*, *Learning*, *Research*, and *Students* were the high frequent words after the function words (that encompassed the first 15 place in the corpus and exceed more than 30% of the data, except the rest of the function words) such as 'the' and its token was 512 (5.4%) from the total token of 9416. Similar to soft science, in hard science, the rate of function words, which incorporated the 31 first place and their rates exceed more than 34000 (11%) from the total token of 311550. The high frequent content words were the five words of *Patients*, *Studies*, *Treatment*, *Data*, and *Method* with the total rate of 27170. For the aim of the current study, the researchers chose two words from each discipline, one specific word, and the other the general in order to explore the collocates of the words. Table 4 shows the collocates of the word Language along with the rates of collocates in left and right sides.

Table 4

The Top Collocates of 'Language' in the Soft Sciences Corpus

Language	Token/2164	Left	Right
Second	407	380	27
Learning	296	54	243
Foreign	194	97	97
English	191	158	33
Learners	170	45	125

Table 4 indicates the top collocates of 'Language' occurring immediately prior and subsequent to the node word in the soft science. The notable frequency of “language” in soft science makes it evident that the authors of soft discipline are mainly interested in studying on the concept of language in both majors of applied linguistics and psychology that the data gathered of them. This idea becomes more evident when other content words referring to “language” appear with high Keynes such as “learning, teaching, students, education, knowledge”. Based on the findings obtained from data analysis, the first high frequent collocate of *Language* is *second* with the token of 407 from the total token of 2164 that it tended to occur more left position than the right one. It needs to be noted that this term has a larger predisposition for the left hand side location of the word 'language,' as it appears here more frequently than on the right hand side. Furthermore, the primary landing site for 'language' was second, with 407 instances, compared to only 27 in Right place. This means that when using 'second' with 'language,' one is considerably more likely to utilize it in the left position rather than the right. The fourth collocate of 'English' was the same. The third most commonly used collocate, 'learning,' tended to appear fairly evenly on two directions of the node word, with 97 and 97 instances, accordingly. The lowest occurrence belonged to the word 'learners', which its right position weighted more than the left one. Table 5 shows the collocates of the word research as a general word in all disciplines, no difference in soft or hard, along with the rates of collocates in left and right sides.

Table 5

The Top Collocates of 'Research' in the Soft Sciences Corpus

Research	Token/13100	Left	Right
Writing	123	75	48
Language	118	74	44
Study	74	34	40
Previous	67	65	2
Second	58	39	19

Table 5 shows the five top collocates of 'research' occurring immediately prior and subsequent to the node word in the soft science. The first high frequent collocate of *research* is *writing* with the token of 123 from the total token of 21640 that it tended to occur more in left position than the right one. The second, the fourth, and the fifth collocates *(language, previous*, and *second*) tended to occur more in left hand than the right side position, however, the collocate of study in the third raw tended to occur with 40 occurrences in right rather than the left. The above procedure was continued for hard science. Finally, Table 6 shows the collocates of the word 'patient' as a specific word in hard discipline, along with the rates of collocates in left and right sides.

Table 6

The Top Collocates of 'Patient' in the Hard Sciences Corpus

Patient	Token/1560	Left	Right
Treatment	12	6	6
Safety	10	2	8
Clinical	9	5	4
Population	8	0	8
Radiation	7	4	3

As Table 6 demonstrates, the five collocates of the word *'patient'* in the hard sciences corpus were content words as well. The first high frequent collocate of *patient* is *treatment* with the token of 12 from the total token of 1560 that it tended to occur equally in both left and hard positions. The second collocate that is 'safety' has a stronger tendency for the right hand side position. Moreover, the third collocate of *clinical* was the main landing site for *'patient'*, with 9 occurrences, whereas it only occurred 5 times in left position and 4 times in right position. The same was with the fifth collocate of 'radiation', whereas in the collocate of 'population', the left position was empty and all of 8 tokens were placed in right hand.

Regarding semantic prosody, the examples of both corpora in terms of general collocations are presented below. Two terms of *research* and *studies* were used in both disciplines, so they were suitable for evaluative prosody analysis. Example 1 shows one instance of the term *research* along its collocates from soft science and example 2 from hard science.

Example 1: "*More research* is needed on different reading related motivational aspects, such as task values (Baker & Wigfield, 1999; Chapman & Tunmer, 2003; Wolters, Denton, York, & Francis, 2014), and the emotions associated with reading experiences (Efklides & Volet, 2005; Pekrun & Linnenbrink-Garcia, 2012)" (Päivinen et al., 2018, p. 1724).

The collocation of *research* with the term *more* in this example mostly appears in the main clause, the study of its evaluative prosody too is focused on the main clause. Nevertheless, as previously noted, determining whether an item's evaluative prosody is positive, neutral, or negative is individual, even natural. Based on this, this collocation is negative; the collocates then function to introduce the main problem existed in the lack of studies on different reading related motivational aspects and the emotions associated with reading experiences. In fact, here, the collocate of *more research* is negative as it refers to the lack of the studies in the field. The same term with the collocate of *recent* is obtained from hard science (example 2), surely with different semantic prosody.

Example 2: "*Recent research* into the cellular and molecular events underlying the development and progression of atherosclerosis has shown that atherosclerosis is a dynamic, inflammatory process" (Weissberg, 2012, p.247).

In example 2, the collocate of recent research was used as positive semantic prosody since it resulted in a positive factor and step (atherosclerosis is an inflammatory and dynamic process) in solving a vague issue in the "cellular and molecular events underlying the development and progression of atherosclerosis" that was "atherosclerosis is an inflammatory process, dynamic" (Weissberg, 2012, p. 247), as a matter of fact, before recent studies, the researchers were unaware of the a dynamic, inflammatory process of atherosclerosis. Example 3 is one example from the general word of *research* that was used in soft discipline and shows the other aspect of semantic prosody that is neutral one.

Example 3: "Our interest in exploring Morena and Diego's writing experiences is partly motivated by our observation that *research on doctoral writing* tends to treat EAL writers as a homogenous group, paying little attention to candidates from non-Asian countries" (de Magalhães, Cotterall, & Mideros, 2019, p.2).

The collocate of *research on doctoral writing* depicts the neutral evaluative prosody by itself, but in general it has the negative meaning due to the last part of sentence that is "paying little attention to candidates from non-Asian countries" (de Magalhães, Cotterall, & Mideros, 2019, p.2). The other general term found in both corpora with high frequency is the term *method.* Example 4 is an instance of this term along with its collocate obtained from hard science.

Example 4: Insertion of self-expanding metallic stents is now a *well-established method* to treat acute obstructions of the colon and rectum.

As it is axiomatic, the collocate of *well-established method* contains a positive meaning based on the information gathered from the previous part of the sentence that is 'Insertion of self-expanding metallic stents' and resulted in 'treating acute obstructions of the colon and rectum'. The other example from soft science regarding *method* and its collocates is presented in example 5 below:

Example 5: "Within *task-based research*, this study is methodologically innovative in that it employed a combination of *research methods*, including behavioral measures of online keystroke logging and introspective data obtained through stimulated recall" (Révész, Kourtali, & Mazgutova, 2017, pp.210-211).

Above two terms of *research* and *method* as general nouns are collocated with task-based and research. The first one that is task-based research is a neutral one in terms of prosody since there are lots of research fields in the world that they have no priority to each other, rather they complement each other, but the same is not true about *research methods* as the study is methodologically innovative since the combination of *research methods* made it innovate, hence it created a positive sense in the sentence. Without doubt, as the findings of data analysis showed the context of text and the co-text play critical roles in making the collocations as neutral, positive and negative, regardless of the subjective nature of evaluative prosody. The results showed that the discipline in general terms such as research and method did not significant effect on semantic prosody and its different senses, but surely it affects the specific terms. Consider the terms *patient* at example 6 obtained from hard discipline:

Example 6: “Venous access is frequently needed in *cancer patients*, especially for intermittent chemotherapy infusions” (Zaghal, 2012, p.207).

The word *'cancer patients*' exposes a negative meaning on the reader, undoubtedly, and this sense is the same in example 7 (aphasic patients) with a different discipline that is soft discipline.

Example 7: “To this end, we present actual data from 59 *aphasic patients* on naming and word repetition tasks, and compare the effect of word frequency on their performance on the two tasks” (Nozari et al., 2010, 542).

The whole sentence is informative in nature, but the collocate *aphasic patients* has a negative meaning and sense in general. The same can be claimed with *'intelligent students' learning outcomes'* that the term intelligent has a positive meaning in itself and consequently, the collocate of *'intelligent students'* did not change the sense of meaning.

Discussion

As stated, the main aim of the current research was to explore the rates of collocations used in the introduction sections of hard and soft disciplines. Furthermore, the evaluative prosody of general collocates in each disciplines were reviewed. The findings revealed that the frequency of function words such as articles were high in frequency, which in the current study, they were ignored and the frequencies of content words, including specific and general words, were estimated. Based on the results, in soft science, the five content words of *Language*, *Writing*, *Learning*, *Research*, and *Students* were the high frequent words after the function words with the total token of 9416. For instance, *language* as the most often used node word in soft science collocated with the common collocates such as *ability, second, foreign, art, assessment, association, background, cognition, community, context, deficit,* and *development*, mostly in right hand rather than left hand. Similar to soft science, in hard science, the high frequent content words were the five terms of *Patients, Studies, Treatment, Data,* and *Method* with the total rate of 27170. Phrases that use the word node of *patient* in hard science (as the high frequently used node word) and its lexemes include, *patients affected by, patients compared to/with, patients receiving (+N) , patients suffering from,* and *patients treated by.* At these phrases, the node of *patient* was placed at left hand and it was collocated with the verbs and prepositions such as *suffered, compared, by,* and *to*.

However, the node word of *patient* collocated with the words such as *cancer, trauma, HCC, MDS, PTSD,* and *young,* in the right hand. Moreover, as it is clear crystal, some of the collocations like *young* have positive meaning in terms of evaluative prosody, however, collocating with the word node of *patient,* the *young patient* changes its semantic prosody and takes a negative sense. In fact, context and co-text influence the semantic prosody of collocations. Also, based on the findings, the rate of collocates in hard science was more than that of soft science. The results of this study provide strong assistance to the theory of lexical priming put forward by Hoey (2005). A term is collectively filled with the contexts and co-texts where it is presented, as per this theory, and our interpretation of it involves the fact that it co-occurs in certain kinds of context with certain other words. In the sense of its incidence, a term is thus primed for its collocates and for the locations of those collocates.

In technical terms rather than in general words, the power of co-occurrences of collocates was high. The same still refers to the incidence locations of those identical collocates that held various positions in our corpora. Knowledge of general and specific collocations is deemed a compulsory element of academic

reading and writing skills that can be closely attributed to the job opportunities, academic success, economic well-being and public status of individuals (Selmistraitis, 2020). Inadequate awareness of collocation causes problems in constructing academic texts through fields at various levels of education. The study results can allow those writing scientific research papers and academic texts to recognize slight changes in the use of collocations in subject areas. It will also allow researchers to use naturally occurring concordance lines from authentic texts relying on corpus-based studies to enhance understanding of collocations in the English language and establish natural language stream in their scholarly papers.

The other focus of the study was semantic prosody across disciplines. The results revealed the positive, neutral, and negative senses of the general collocations regardless of their disciplines. It means that the collocates themselves expressed the sense and being in soft or hard science did not affected their semantic prosody, for instance that collocates such as *mixed-method, quantitative method* and *qualitative method* are used in two disciplines and they are neutral in terms of semantic prosody, hence the context (soft and hard) does not influence the sense. This finding suggests that, given diverse disciplines, the authors' perceptions and awareness of semantic prosody may be similar. Instead, the main distinction is between word collocation and their generality and specificity. As a result, it is proposed that while semantic prosody perhaps not always cause problematic issues in writing for authors in soft and hard fields, collocations may nevertheless be a difficult nut to crack for them. Surprisingly, Bi's analysis yielded the same result (2019). By comparing Chinese students' writing with English natives' writing, he explored the applications use of frequency, collocation, and semantic prosody of two sets of synonyms. While the collocational pattern varies significantly between the two, the semantic prosody is the same.

The findings are consistent with previous research on semantic prosody conducted by Louw (1993), Sinclair (2004), and Hunston (2011); nevertheless, these analyses have indicated that the affective meaning of a term can only be understood while it is used in the context of its routine collocations. The reason for the difference was the fact that they studied specific words of each discipline, whereas, in the current study, the semantic prosody of general words was analyzed and the general and specific words were analyzed in terms of collocations.

The results are in congruent with Shin's (2019) study that like the current study they reviewed collocations and semantic prosody, but in the academic writing of first-year students at university level who were native and nonnative English. Semantic traits unique to each language group are suggested by the findings. In addition, the results of this study are in congruent with Stewart's study in terms of collocations and evaluative prosody (2010). The researcher studied the *break-out* verb, and he identified 1,126 instances of the verb in the BNCC. The data outcome demonstrated that, *break out* demonstrated a semantic preference for 'conflict situations,' 'illness,' or, more generally, 'difficult situations,' because the following words are found in the current context of break out:

infection, war, crisis, conflict. The verb is linked with an unfavorable sense of meaning or semantic prosody, which is dependent on its semantic preferences, because it cannot be characterized as an element whose underlying meaning is unpleasant (Stewart, 2010). Begagi (2013) looked at semantic prosody and semantic preference of one of the most prevalent verb-noun collocations in the Corpus of Contemporary American English, which was relevant to this research (COCA). The researchers noted that all word variants of the collocation make sense more often in a pessimistic journalistic record than in a scholarly one. In general, the authors found that it cannot be confirmed that the collocation has an extremely negative prosody, since it has been shown that in a negative environment some word forms of meaning occur more

frequently than other word forms, while some others occur more frequently in a favorable manner. The findings of the current study will contribute to education in argumentative essays, that are a popular type of pedagogical writing allocated to undergraduate and language-focused programs across disciplines.

We may draw a variety of conclusions taking into account the results of the distinctions made between how terms are understood in hard and soft sciences with respect to the collocates. Firstly, it is shown that a phrase, such as 'language' or 'patient', appears to combine with some distinct and specific terms in general in different disciplines, and some distinct and special functions and content. Second, although they were overlooked in the current study, the frequencies of feature words are high in two disciplines; most of the content words, nevertheless, are not comparable in prevalence throughout different disciplines. In addition, this analysis concludes that the general words collocate themselves directly, without the interference of the context, expressed the semantic prosody and the meaning they transferred to the reader. It is worthwhile saying that since the data for the current research was gathered from ISI research articles as notable ones, hence reporting the frequencies of different categories of lexical priming such as collocations can have pedagogical implications for novice authors in both hard and soft science to pay attention to the high frequent categories and types (for example one special type of collocation that highly used by authors of RAs) and used them as models in their writings as a pattern. Given that the major problem in learning may lay in word collocation rather than semantic prosody awareness, English instructors are advised to focus their efforts on assisting L2 students with word collocation rather than semantic prosody awareness. Instructors could use corpus-based research like the current one to enlighten students about the distinctions in collocations in terms of the majors they are studying to help them grasp the different uses of words, improve their own vocabulary, and therefore raise L2 knowledge since as Wilkins (1972) believed little can be communicated without grammar, and nothing can be communicated without vocabulary. Also, unawareness of the nature of collocations, in terms of general and specific collocations related to each science, perhaps compensate for the abuse, overuses, and underuses of English collocations to a significant extent, hence studies like the present one can assist learners and writers of research articles to scrutinize on the nature of collocations and the sense that collocations create in the form of semantic prosody. The same study can be conducted with soft and high science paying attention to the colligations and semantic preferences. Moreover, the culture was ignored in the current study that the interested researchers can scrutinize on this issue.

References

Anthony, L. (2011). Antconc (Version 3.2. 4w) [Computer Software]. *Tokyo: Waseda University, [cit. 31.5. 2015]. Dostupné na: http://www. antlab. sci. waseda. ac. jp/antconc_index. html*.

Baker, P., Hardie, A., & McEnery, A. (2006). *A glossary of corpus linguistics*. Edinburgh, Scotland: Edinburgh University Press.

Begagić, M. (2013). Semantic preference and semantic prosody of the collocation make sense. *Jezikoslovlje*, *14*(2-3), 403-416.

Bi, Z. (2019). A Semantic Prosody Analysis on Two Synonymous Pairs in English Native Speakers' and Chinese Learners' Writings. *English Language Teaching*, *12*(8), 14-19.

Bublitz, W. (1996). Semantic prosody and cohesive company: Somewhat predictable. *Leuvense Bijdragen: Tijdschrift voor Germaanse Filologie,85*(1– 2), 1-32.

Conklin, K., & Schmitt, N. (2008). Formulaic sequence: Are they processed more quickly than non-formulaic language by native and nonnative speakers? *Applied Linguistics, 29*(1), 72-89.

Dam-Jensen, H., & Zethsen, K. K. (2008). Translator awareness of semantic prosodies. *Target,*

20(2), 203–21.

Davis, M., & Morley, J. (2015). Phrasal Intertextuality: The Responses of Academics from Different Disciplines to Students' Re-Use of Phrases. *Journal of Second Language Writing*. *28*(2015), 20–3.

Du, Y., Wang, B., & Yang, L. (2018). A Corpus-Based Study on the Semantic Prosody of Quasi-Affix "Zu". In *Workshop on Chinese Lexical Semantics* (pp. 445-453). Springer, Cham.

Firth, J. R. (1957). *Papers in linguistics 1934–1951*. London, England: Oxford University Press.

Firth, J.R. (1968). A synopsis of linguistic theory 1930–1955' in F. R. Palmer (ed.): *Selected Papers of J. R. Firth 1952–59*. Bloomington: Indiana University Press, pp. 1–32.

Gilquin, & Paquot, M. (2008). Too chatty. Learner academic writing and register variation. *English Text Construction, 3*(1), 41-61.

Greenbaum, S. (1974). Some verb-intensifier collocations in American and British English,' *American Speech, 49*, 79–89.

Halliday, M. A. K. (1966). Lexis as a linguistic level' in C. Bazell, J. Catford, M. A. K. Halliday, and R. Robins (eds): *In Memory of J.R. Firth*. London: Longman, pp. 148–62.

Hoey, M. (1991). *Patterns of Lexis in Text.* Oxford: Oxford University Press.

Hoey, M. (2003). Why grammar is beyond belief. In J.-P. van Noppen, C. Den Tandt, & I. Tudor (Eds.), *Beyond: New perspectives in language, literature and ELT* (*Belgian journal of English language and literature, new series, 1*, pp. 183–96). Ghent, Belgium: Academia Press.

Hoey, M. (2004). The textual priming of lexis. In S. Bernadini, G. Aston, & D. Stewart (Eds.), *Corpora and language learners* (pp. 21–41). Amsterdam, Netherlands: John Benjamins.

Hoey, M. (2005). *Lexical priming: A new theory of words and Language.* London: Routledge.

Howarth, P. (1998). The Phraseology of Learners' Academic Writing. In A. Cowie (Ed.), *Phraseology: Theory, Analysis and Applications* (pp.161– 186). Oxford: OUP.

Hu, H. C. (2015). A Semantic Prosody Analysis of Three Adjective Synonymous Pairs in COCA. *Journal of Language and Linguistic Studies*, *11*(2), 117-131.

Hunston, S. (2007). Semantic prosody revisited. *International Journal of Corpus Linguistics, 12*(2), 249-268.

Hunston, S., & Thompson, G. (Eds.). (2000). *Evaluation in text: Authorial stance and the construction of discourse: Authorial stance and the construction of discourse*. Oxford University Press, UK.

Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of john Sinclair* (pp. 157-176). Amsterdam: John Benjamins.

Louw, B. (2000). Contextual prosodic theory: Bringing semantic prosodies to life. In C. Heffer, H. Sauntson, & G. Fox (Eds.), *Words in context: A tribute to john Sinclair on his retirement* (pp. 48-94). Birmingham: University of Birmingham.

Oster, U. (2010). Using corpus methodology for semantic and pragmatic analyses: What can corpora tell us about the linguistic expression of emotions? *Cognitive Linguistics, 21*(4), 727-763.

Partington, A. (1998). *Patterns and Meanings*. Amsterdam: John Benjamins.

Partington, A. (2004). Utterly content in each other's company: Semantic prosody and semantic preference. *International Journal of Corpus Linguistics, 9*(1), 131-156.

Krishnamurthy, R. (2000).Collocation: from *silly ass* to lexical sets' in C. Heffer, H. Sauntson, and G. Fox (eds): *Words in Context: A Tribute to John Sinclair on his Retirement*. Birmingham: University of Birmingham.

McEnery, A. & Wilson, A. (2001). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Martín, P., Rey-Rocha, J., Burgess, S., & Moreno, A. I. (2014). Publishing research in English-language journals: Attitudes, strategies and difficulties of multilingual scholars of medicine. *Journal of English for Academic Purposes, 16*, 57-67.

Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics, 32,* 111-129.

Sardinha, T. B. (2000). Lingüística de corpus: histórico e problemática. *Delta: documentação de estudos em lingüística teórica e aplicada, 16*(2), 323-367.

Schmitt, N., & Carter, R. (2004). *Formulaic sequences in action: An introduction. Formulaic sequences.* Amsterdam: Benjamins.

Selmistraitis, L. (2020). Semantic Preference, Prosody and Distribution of Synonymous Adjectives in COCA. *GEMA Online® Journal of Language Studies, 20*(3).

Shin, Y. (2019). Do native writers always have a head start over nonnative writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes, 40*, 1-14.

Sinclair, J. (1991). *Corpus, concordance, collocation: Describing English language*. Oxford: OUP.

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative methods,' *Function of Language, 2*(1), 1–33.

Swales, J. M., & Feak, C. B. (1994). *Academic writing for graduate students* (pp. 155-6). Ann Arbor, MI: University of Michigan Press.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam, Netherlands: John Benjamins.

Tribble, C. (2011). Revisiting apprentice texts: Using lexical bundles to investigate expert and apprentice performance in academic writing. In F. Meunier, S. De Cock, G. Gilquin, & M. Paquot (Eds.), *A taste for corpora. In honour of Sylviane Granger* (pp. 85-108). Amsterdam: John Benjamins.

Wei, Y., & Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELC Journal, 42*(2), 155-166.

Wilkins, D. A. (1972). *Linguistics and Language Teaching*. London: Edward Arnold.

Wilson, J. L. (2012). *Carbonate facies in geologic history*. Springer Science & Business Media.

Xiao, R., & Mcenery, T. (2006). Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. *Applied Linguistics, 27*(1), 103-129.

Biodata

Leila Ghaemi is a Ph.D. candidate at Tabriz Islamic Azad University. She has been researching different areas of teaching and he is interested in discourse analysis and corpus-based studies.

Davud Kuhi is an academic member of the Department of English Language at Islamic Azad University. He researched on different properties of academic discourses for approximately two decades and has published many articles in prestigious national and international academic journals.

Nesa Nabifar is an assistant professor at Tabriz Islamic Azad University. She holds a Ph.D. in linguistics and ELT. She has published several papers at different international and national conferences.