



mph.sdcongress.ir

هشتمین همایش بین المللی
مدیریت، روانشناسی و علوم انسانی
با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



The Rhetorical structure of Literature Reviews in English and Persian Research Articles

Nasim Mehrabian

English Department, Najafabad Branch, Islamic Azad University, Najafabad, Iran

E-mail: nas.mehrabian@gmail.com

Abstract — This study reported here focused on identifying linguistic complexity indices in the review sections of English and Persian research articles which can be used to discriminate among various disciplines. Lexical diversity and lexical density as the measures of lexical complexity, as well as average sentence length and the ratio of subordination as two measures of syntactic complexity were used in this study. Another linguistic feature called move complexity was centered around, too. Swale's CARS model was taken into consideration for figuring out the move patterns. Linguistic complexity indices in four major disciplines, namely humanities, engineering, basic sciences, and medical sciences were examined. The corpus comprised of 288 review sections of RAs (144 in English, 144 in Persian). From each discipline, 72 articles were tested. Wordsmith Tools software was utilized to compute the four lexical and syntactic complexity measures of each piece of data. MANOVA and discriminant function analysis were run, as well. The results demonstrated the highest degree of lexical and syntactic complexity among Persian and English articles in the field of humanities. The second significant difference was found among engineering articles in Persian and their English counterparts. The articles of medical sciences and basic sciences showed no difference. This study has implication for ESP and EAP teachers and practitioners.

Keywords-Literature Review; Lexical Complexity; Syntactic Complexity; Move Complexity



mph.sdcongress.ir

هشتمین همایش بین المللی
مدیریت، روانشناسی و علوم انسانی
با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



1. Introduction

Writing a correct literature review is the most imperative step of the analysis method in quantitative, qualitative, and mixed research studies (Boote & Beile, 2005; Combs, Bustamante, & Onwuegbuzie, 2010; Onwuegbuzie, Collins, Leech, Dellinger, & Jiao, 2010). As noted by Boote and Beile (2005), "A complete, advanced literature review is the foundation and inspiration for valuable research. The complicated nature of research requests such exhaustive, thorough, reviews". The literature comprehensively alludes to data significant to the topic of interest. The literature deals with published scholarly works, including journal articles, books, conferences, and technical reports. A literature review audits the academic writing on a particular subject by outlining and examining distributed work on that theme. It portrays, analyzes, differentiates and assesses the main theories, arguments, methodologies, philosophies, approaches, and controversies in the academic literature on a subject matter. It will determine what is identified (and unknown) within the branch of knowledge, acknowledge zones of debate and facilitate formulate questions that require more analysis.

There are various reasons why a review of related literature remains the central part of any logical study. These include; First, review of literature goes about as a stepping-stone towards the accomplishment of the goals of a study. Second, literature reviews provide a strong foundation to back one's investigation. Third, literature reviews assist the scientist with avoiding duplication, recognizing the gaps in different studies with the aim of filling them, borrowing from the research plan and strategy used to investigate that specific issue and to explain the new discoveries.

As the most difficult part of any article, the literature review section merits careful examination. The most specified linguistic features of the literature review section such as lexical complexity, syntactic complexity, and move complexity seem to fail to attract many researchers' attention. Thereupon, the aim of this study was to focus on these aspects of English and Persian reviews of literature and make use of them to differentiate between various kinds of writing in a number of disciplines. Research articles have been subjected to various linguistic features investigation. For instance, Marco (2000) investigated collocational frameworks in medical scientific articles. As another example, use of the directives in hard sciences was researched by Hyland (1999). Hyland (2008) analyzed the lexical bundles and found significant variations across disciplines. As indicated by previous literature, fluency, accuracy, and complexity are among the fundamental measurements used for describing linguistic features of different texts (Pallotti, 2009).

2. Literature Review

2.1. Review on Syntactic Complexity

The primary terms used in this study were lexical complexity, syntactic complexity, and move complexity. Lexical complexity and lexical diversity were presented as the most widely-used complexity indices to describe the productive lexicon (Laufer & Nation, 1995). Different studies on syntactic complexity (cf. e.g. Wolfe-Quintero, Inagaki and Kim 1998, Norris and Ortega 2009) have additionally analyzed the methods for estimating syntactic complexity. As noted by Kuiken and Vedder (2008), lexical density and diversity have a direct relationship with task complexity and difficulty. Lexical diversity estimates the extent of vocabulary in a content or of an author. Instinctively, lexical diversity should be quantifiable as far as various word composes. Since this measure is known to depend heavily on text size, the number of tokens are likewise considered. In



mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



Bachman's (1990) conceptual model of language ability, syntactic complexity turned out to be a compelling component in the second language assessment construct. Ortega (2003, p.492) described syntactic complexity as "the range of forms that surface in dialect production and the level of complexity of such forms". The development of the syntactic complexity of students' writing has been at the focal point of various examinations in recent years (e.g., Byrnes, 2009; Byrnes & Sinicrope, 2008; Crossley & McNamara, 2014; Shaw & Liu, 1998; Vyatkina, 2013), however just a couple of studies thought about lexical improvement related to syntactic changes in students' written production (Bulté & Housen, 2014; Storch & Tapper, 2009; Verspoor, Lowie, & van Dijk, 2008; Vyatkina, 2012). Syntactic complexity, i.e. the range and level of advancement of syntactic structures that surface in language production, has been perceived as an imperative construct in second language writing teaching and research (e.g. Ortega 2003).

According to Li (2000), by linguistic qualities of learners' composition, we mean syntactic complexity, lexical complexity, and grammatical accuracy. Syntactic complexity is assessed by (1) average sentence length, and (2) the ratio of subordinated structures. Average sentence length is acquired by figuring the average number of words per sentence in a given content. The ratio of subordination is measured by calculating the ratio of the number of subordinated structures to the combination of subordinated and coordinated structures in a piece of writing. Lexical complexity is evaluated by (1) lexical diversity, and (2) lexical density. Lexical diversity is used in applied linguistics and related fields to refer to the displayed range of vocabulary in a given text. Having the capacity to show a large vocabulary is frequently viewed as a critical part of being a sophisticated language user. Higher lexical diversity is generally considered to demonstrate further developed capability than lower lexical diversity (Malvern, Richards, Chipere, & Duran, 2004). Lexical diversity is calculated by having the range of various words together with each content and function words divided by full range of words in a piece of writing. Lexical density is calculated by having the number of lexical items excluding function words divided by the total number of words in a piece of writing. There are two alternative ways to step on grammatical accuracy: (1) the magnitude relation of range of grammatical errors to the whole range of sentences, and (2) the magnitude relation of styles of grammatical errors to the whole range of sentences in a piece of writing. Lexical diversity and essay length can precisely predict essay ratings.

With respect to syntactic complexity, Polio (2001) recognized the average length of a structure, the frequency of a structure and complexity ratios as the most widely used strategies for estimating complexity. The number of words per T-unit is used for measuring average length of a structure, and complexity ratios like coordination index are used to measure the ratio of independent clauses per clauses combined. By checking particular structures, such as passive sentences or dependent clauses within a specific time frame the frequency of a structure is examined.

There has been a progression of experimental examinations inspecting the relationship of syntactic complexity measures to L2 proficiency. Therefore, syntactic complexity is the use of varying structures with complex elements such as embedded dependent clauses. The aim of syntactic complexity analysis is to describe such complex structures, i.e. to see how smaller units and basic sentences are joined into more unpredictable structures (Holger 2004: 3). As Hunt (1965) states, more are frequently viewed as an indication of complexity. In this way, length, amount of embedding, and recurrence of certain complex structures can work as a reason for syntactic complexity. Thus, long production units can be considered more complex than short units. The number of subordinated structures is one of the most widely used measures in syntactic complexity studies, particularly in L2 studies (Ellis and Barkhuizen 2005: 154). Holger (2004: 3) states that complex sentences are the result of simple sentences that are connected together,



mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



through coordination and subordination. This connection of production units makes the language more complex. Bergman and Abrahamsson (2004: 611) have made a three-level scale for portraying the syntactic structures in L2. For beginners, the structures of a simple sentence are used and only the basic linking elements (such as and, but, then) are available. At the intermediate level, the number of complex sentence structures increases, particularly dependent clauses, with variation in the use of linking elements. At this level, infinite clauses appear in the production of learners. The advanced learners' level is described by various sentence structures with numerous dependent and infinite clauses. Prior studies regularly centered around the development of syntactic complexity (e.g. Bardovi-Harlig, 1992; Mellow 2006) or the task types' effect on L2 written (e.g. Ishikawa 1995, 2007; Storch and Wigglesworth 2007; Kuiken and Vedder 2007; Robinson 2007) and spoken complexity (e.g. Skehan and Foster 2005; Tavakoli and Foster 2008). Differentiating spoken and written material was also another way for researchers to study syntactic complexity. As an example, in the 1960s, the main focus of studies comparing spoken and written language was on word frequency counts (related to the length of sentences) and brought about finding that writing was a lot of advanced than speech.

2.2. Review on Lexical Complexity

Lexical knowledge is likewise one of the fundamental requirements for academic accomplishment. Engber (1995) noticed that the length of a paper will influence the lexical variety: "the longer essays are more likely to have repeated lexical items". Engber (1995) likewise found that there is a considerable correlation between lexical diversity and holistic quality ratings of the compositions. Engber (1995), Laufer and Nation (1995), (1981), McNamara et al. Mellor (2011), Grobe (2010), and Yu (2010), expressed that the correlation between lexical diversity and overall quality ratings of writings was statistically vital and positive. Higher correlations ($r = 0.7$ and $r = 0.6$) were found by Koda (1993) between lexical density and therefore the quality of writings and he introduced lexical diversity jointly study predictor of writing quality. Berman (2000) led a far-reaching investigation of lexical diversity among seven languages, compared young people at three ages, plus adults, produced language in two genres of narrative and expository, and two modalities. He discovered principle impacts for language, genre, and age, but not for modality. The results of Woerfel and Yilmaz (2011) study which demonstrated that age and genre may modify the measurement of lexical diversity, word and text length, supported the results of Berman's investigation. According to these two studies, lexical diversity differs among different genres. In order to measure lexical features, Polio (2001) noted that lexical features frequently mean lexical richness and that they nearly manage the size and variety of lexicons. These features may not be measured simply by counting the total number of words or by using a type-and-token ratio. The topic and the task will also decide the scope of vocabulary used in an essay (Polio, 2001).

Great advances have been made in measuring lexical diversity in oral or written speech samples, and more sophisticated indices have been created, despite the fact that some researchers (e.g. Le Normand, Parris, & Cohen, 2008) are still measuring lexical diversity by means of counting the number of different words or the total number of words in a text. As a general purpose measure of spoken and written language development, lexical diversity measures have been considered (Malvern 2004). Ong and Zhang (2010) used two measures of type-token ratio and an improved type-token ratio to investigate the effects of cognitive task complexity on the lexical complexity in writing. For the first measure the length of the text is not important, but in

the second measure, the number of word types per square root of two times divided by the total number of word tokens was calculated (Wolfe-Quintero et al., 1998).

2.3. Review on Move Complexity

Swales (1990) was the pioneer of move analysis. A move, according to Swales (2004) is a “discoursal or rhetorical unit that performs a coherent informative task in a written or spoken debate”. Swales (1981, 1990) proposed and developed the concept of a move to analyze a textual structure. In 1981, he constructed a four move model of introductions, which included the following: setting up the exploration field, detailing past research, planning and finally presenting the present research. Later he revisited the model. His CARS (Create A Research Space) model has three Moves; each move has some related steps. Move 1, establishing territory (setting up an area), is an explanatory instrument used by the academicians to begin their study. The steps of M1 are: claiming centrality (M1S1), topic generalization (M1S2), and reviewing the related studies (M1S3) which consistently occur at the beginning of the research. The second move is establishing a niche (M2) by counterclaiming (M2S1), indicating a gap (M2S2), question-raising (M2S3), and continuing the tradition (M2S4) which separately contend on the misinforming of the past examinations, show the inadequacy of the previous investigations, question the past discoveries, and guarantee that new clarification for the discoveries is required. The third Move of the CARS model is occupying the niche, in this Move, writers represent the located gap in the second Move. This Move can be understood by outlining purposes (M3S1), reporting present research (M3S2), declaring main findings (M3S3), and showing RA structure (M3S4).

Numerous studies have been done on different sections of research articles regarding Move analysis, e.g. The Introduction section (Bhatia, 1993; Ozturk,2007; Monreal,2008), Abstract section (Hyland ,2000; Samraj, 2005), Discussion section (Fallahi and Erzi, 2003), and Result section (Brett,1994).

A literature review can be process and product. To Prepare a literature review, some steps must be taken including; defining the subject matter and the scope of the review, searching the library catalogs, subject-specific databases and other search tools to discover sources that are applicable to the point, reading and evaluating the sources and to decide their appropriateness to the understanding of the topic at hand, and analyzing, interpreting and discussing the findings and conclusions of the sources. Thus, in this study, the purpose was to make a correlation between the literature review sections of both English and Persian research articles written in four disciplines namely; humanities, basic sciences, medical sciences, and engineering to distinguish the similarities or differences in terms of lexical complexity, syntactic complexity, and move complexity. Few studies have already investigated these three linguistic features simultaneously in English and Persian research articles.

The present investigation along these lines endeavored to make a commitment to learning by analyzing articles in their reviews, describing the rhetorical structure of research articles and variations within a discipline. Thus, the key aim of the present study was to examine the syntactic, lexical and move complexity in the review of the literature of English and Persian published research articles. For this research four disciplines were chosen including humanities, medical sciences, basic sciences, and engineering. The present study used Swales’ theoretical framework (1990; 2004) to distinguish moves and steps type / token ratio for examining lexical complexity, average sentence length and the ratio of subordination to recognize syntactic complexity based primarily on the content of each article.



وزارت آموزش عالی و تحقیقات علمی

mph.sdcongress.ir

هشتمین همایش بین المللی
مدیریت، روانشناسی و علوم انسانی
با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



This study was an attempt to find appropriate answers to the following questions:

1. To what extent, the lexical, syntactic and move complexity in the English and Persian research articles are similar or different?
2. Is there any relationship between the Fields of Study and the ways Lexical, Syntactic and Organizational features of the literature review sections are realized?

3. Methodology

The methodology of the present investigation was established in corpus linguistics. Corpus linguistics is the study of language as expressed in corpora or samples of the texts. Corpora now are automatically derived from the source texts. Corpus linguistics proposes that reliable language analysis is more possible with corpora collected in the field, in their natural context, and with least experimental inferences. Corpus linguistics collects and stores, authentic written and spoken information, that is used as the source of data for text analysis so as get sufficient information about texts. The great advantage of corpus- linguistic method is that language researchers can draw on a large amount of authentic, naturally occurring language data produced by a variety of speakers or writers in order to confirm or refute their own hypotheses about specific language features on the basis of an empirical foundation.

3.1. Corpora

The corpora of the study comprised of 288 research articles (144 in Persian, 144 in English) from four fields of study namely: humanities, basic sciences, medical sciences, and engineering. For each field of study, three subfields were incorporated. Humanities' subfields were management, psychology, and sociology. For basic sciences, chemistry, physics, and mathematics, for engineering mechanics, civil engineering, IT and for medical sciences, nursing, medicine, obstetrics, and gynecology were the subfields. 72 articles of each discipline and 24 papers from every subfield of the study were examined. Humanistic articles were published in Iranian Journal of Applied Linguistics, Journal of Humanities, Iranian Journal of Economic Studies, International Journal of Management, International Journal of Psychology, and International Journal of Sociology. The articles of basic sciences were selected from Iranian Journal of Mathematical Sciences and Informatics, International Journal of Chemistry, Journal of Physics, Journal of Applied Chemical Research, International Journal of Industrial Mathematics, and Journal of Mathematical Chemistry. International Journal of Engineering, International Journal of Advanced designs and manufacturing Technology, International Journal of Mechanics, Iranian Journal of Chemical Engineering, Journal of Civil Engineering, and Journal of Computer Science and Engineering were used as the source of data for the engineering discipline. The journals of medical science were International Journal of Medical Studies, American Journal of Nursing, American Journal of Medicine, International Journal of Gynecology and Obstetrics, International Journal of Nursing, and International Journal of Medical Studies. The majority of the articles were published during the period of 1998-2010 in 24 Iranian and International journals.

3.2. Instruments and Procedures

The purpose of the study was to analyze three linguistic features of review sections of published research articles, namely lexical complexity, syntactic complexity, and move complexity. MTLTD



mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



(a measure of textual lexical diversity) was used for assessing lexical diversity (McCarthy & Jarvis, 2010). This measure is similar to the type-token ratio (TTR), but the TTR is firmly influenced by text length and was therefore not appropriate to compare the texts of different lengths. To overcome this weakness of TTR, MTLTD was developed. For lengthy texts, MTLTD was used. The lexical word frequency measure was accustomed to assess lexical sophistication. It essentially indicated the common appearance frequency of words found within the review sections. Lexical complexity and lexical diversity were introduced as the most widely-used complexity indices to describe the productive lexicon (Laufer & Nation, 1995). Lexical diversity computed by having the number of different words including both content and function words divided by the total number of words in a piece of writing. Among several measures to assess the lexical diversity of texts, type/ token ratio (TTR) was the oldest and most frequently used measure. It provided that lexical diversity could be calculated through the ratio of different words (types) to the total number of words (tokens) (Laufer & Nation, 1995; Johansson, 2008).

Various measures had been proposed for assessing syntactic complexity in the second language writing development literature. Most of these measures checked syntactic complexity by evaluating one of the following: length of the production unit, amount of subordination or embedding, amount of coordination, the range of syntactic structures, and degree of phrasal sophistication. The average length of a structure was measured by words per T-unit technique, and complexity ratios such as coordination index were used to gauge the ratio of independent clauses per clauses combined. Another index of syntactic complexity was the ratio of subordination which has been the main concern of different research studies (Wolfe-Quintero et al., 1998). La Brant (1934) calculated the ratio of subordination by dividing the number of subordinate clauses by the total number of clauses. Average sentence length was also used to measure grammatical complexity because Ortega (2003) in his studies on syntactic complexity in L2 writing, found that grammatical complexity differed greatly in accordance with the L2 proficiency level. It also worked comparatively well with the texts of basic EFL writers (Nitta & Baba). This study drew upon Swales' move framework. Each rhetorical move consists of smaller rhetorical elements referred to as 'steps' by Swales (1990) or 'strategies' by Bhatia (1994).

The initial step of analyzing the review sections of research articles was to convert them into the text-only format. After that, some computerized programs and software were used to measure the lexical complexity and syntactic complexity of the text. L2 Syntactic Complexity Analyzer and Lexical Complexity Analyzer and Wordsmith Tools software estimated the text in terms of complexity. The average number of words per sentence, the ratio of subordination, and the average number of sentences were considered in syntactic complexity. Wordlist which is one of the programs of Wordsmith Tools was used for the analysis of syntactic complexity in terms of average sentence length of a given text. The analysis of move complexity was based on Swales' CARS Model.

3.3. Data analysis

The criteria for selecting the material was MSRT indexed journals (Elmi-pazhuhishi) in four fields of study. The present study analyzed three linguistic features to address the research questions. These features were lexical, syntactic, and move complexity. There were different techniques to address the lexical and syntactic complexity of a text in different studies. Wordsmith Tools (Scott, 2009), a text analysis software, was used to calculate complexity indices for each review section. Lexical diversity and lexical density, as two criteria of lexical complexity, and also



mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on Management, Psychology & Humanities with sustainable development approach



average sentence length and ratio of subordinated structures, as two measures of syntactic complexity, were the linguistic metrics which were used in this study to differentiate among various disciplines. Two computerized programs were used for the text analysis like L2 Syntactic Complexity Analyzer and Lexical Complexity Analyzer designed by Lu (2010; 2012) and Wordsmith tools. Wordsmith Tools is an integrated package of text analysis programs designed to examine the behavior of words in texts (Li, 2000). Wordsmith Tools have three major programs, i.e. Wordlist, Concord, and Keywords. In the present study, Wordlist and Concord were used for the analysis. Wordlist analyzed syntactic complexity regarding the average sentence length of a given text. Likewise, Wordlist could analyze lexical diversity and lexical density because of its statistic function in providing statistics of type/ token ratio and prepared data for sentence length. As a secondary measure of syntactic complexity, the ratio of subordination was calculated by the Concord program through searching for the occurrences of subordinated and coordinated structures in a given text. In this way, this software can tally the number of functional, subordinate, and coordinate structures and furthermore average sentence length and the total number of words of each composition. The analyzer was expected to automate syntactic and lexical complexity analysis of review sections of research articles. For analyzing the moves, range, amount, organization, and linguistic features in each move were taken into consideration. After calculating the linguistic indices for each review section, using multivariate analysis of variance (MANOVA), the differences among four disciplines with reference to their linguistic properties were examined. As the following step, a discriminate function analysis was run to evaluate text complexity metrics which best discriminated among the disciplines.

4. Results

In this study, three linguistic features of a text were taken into account, including lexical complexity, syntactic complexity, and move complexity. Among these variables, lexical and syntactic complexity were considered as dependent variables and move complexity as an independent one. Four major disciplines were chosen for this research including; humanities, basic sciences, engineering, and medical sciences. Various disciplines displayed mean differences regarding linguistic indices to evaluate the complexity of the review sections of the research articles. The equality of group means was estimated to verify the obtained mean differences as significant.

Table 1. Tests of equality of group means

	Wilk's lambda	F	df1	df2	Sig.
Average sentence length	1.912	4.94	4	234	0.172
Ratio of subordination	1.28	76.672	4	234	0.000
Lexical diversity	1.20	75.948	4	234	0.000
Lexical density	0.698	218.126	4	234	0.000

Significant differences among disciplinary groups on each of the linguistic indices were revealed in measuring group means and MANOVA. Table 2 showing group statistics, provided information to predict group membership.

Table 2. Group statistics: Means (standard deviation), F –value

Disciplines	Humanities	Engineering	Medical Sciences	Basic Sciences	F
Average sentence length	38.52(8)	36.2(5.06)	35.18(7.56)	32.6(4.76)	3.20



mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



Ratio of subordination	1.18(0.26)	0.84(1.2)	0.84(0.2)	0.66(0.76)	81.20
Lexical diversity	99.71(19.3)	94.42(6.2)	94.88(9.76)	80.48(7.1)	72.84
Lexical density	150.6(4.32)	137.92(3.26)	139.08(4.96)	117(4.90)	68.46

The inter-correlation coefficients for linguistic variation were calculated to ensure that the linguistic indices selected to use in this study, measure separate linguistic concepts. Table 3 shows pooled within-group matrices which revealed the low correlation between different complexity indices. The low inter-correlation can be proof of the fact that these measures do not assess the similar concept.

Table 3. Pooled within-groups matrices (correlation)

	1	2	3	4
1. Average sentence length	1.00			
2. Ratio of subordination	0.019	1.00		
3. Lexical diversity	0.169	0.246	1.00	
4. Lexical density	-0.036	0.094	0.128	1

On the following stage, a discriminant function analysis was run, in which the four measures of the eigenvalues, the Wilk's Lambda, the classification results, and the discriminant function coefficients were taken into consideration. The number of discriminant functions is constantly one less than the number of groups. Here the focus was on two discriminant functions. The first function is the one which serves to show the maximum difference among groups. The second function indicates the larger proportion of the unexplained variation. Table 4 shows the number and important indices of the discriminant functions. It is known as the table of eigenvalues. It shows that there are two functions that are indicative of the four major disciplines of the study. It also provides information on the relative importance of each function.

Table 4. Eigenvalues

Function	Eigenvalues	% Of variance	%Cumulative	Canonical correlation
1	5.028	180.4	180.4	1.692
2	0.074	0.6	190.0	1.578

Wilk's lambda assessed the significance of the two functions. According to Table 5, Wilk's lambda ratios were significant for the first function ($\chi^2(7) = 138.03, P < 0.001$). This function is highly significant for the study because the second function was considered as non-significant.

Table 5. Wilk's lambda

Test of function(s)	Wilk's lambda	Chi-square	df	Sig.
1 through 2	0.564	292.02	7	0.00
2	1.986	1.688	2	1.678

A correct classification table exhibits the precision of the model used to group the articles into disciplines based on their linguistic features. Perfect predictions require that all cases sit on the diagonal. By means of the ratio of the total number of diagonal cases to the total number of cases, the percentage of correct classification becomes obvious. The review section of the articles about humanities field demonstrated the best classification results (89%). That is, 89% of humanities, 85% of engineering, 50% of basic sciences, 55% of medical science articles were effectively allocated to their disciplines based on their complexity indices. In fact, two of the three linguistic features demonstrated significant differences among various disciplines; ratio of subordination and lexical



وزارت آموزش عالی و تحقیقات علمی

mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



complexity. The highest contribution to group discriminability belonged to lexical density followed by the ratio of subordination. The last contributing factor was lexical diversity.

Humanities articles proved higher linguistic indices than the other three disciplines in their review sections. As the principal linguistic feature, the lexical complexity of review sections of both English and Persian RAs was considered. Its two indices were lexical density and lexical diversity. Lexical diversity concentrated on the number of types and tokens, but lexical density centered on the number of content and function words. After the articles of humanities, engineering, medical sciences, and then basic sciences RAs showed the high linguistic indices. As the following stage, the second linguistic feature, syntactic complexity, was measured. Two indices in particular average sentence length and the ratio of subordination were used. The ratio of subordination is a syntactic complexity index which is computed by having the proportion of subordinate or embedded structures to the combination of subordinate and coordinate structures. The number of subordinate and coordinate clauses as well as the number of different types of sentences, including simple, compound, complex, and complex-compound sentences were assessed. The results of the syntactic complexity of RAs were similar to lexical complexity. The sentences of humanities discipline RAs were lengthier (on average 20 words per sentence) than the other three disciplines. The majority of the sentences of these articles were simple and complex sentences. There were 30 simple sentences and generally equal number of complex sentences. The ratio of subordination of humanities articles was higher than engineering, medical and basic sciences. In fact, humanities articles were more complex in terms of lexical and syntactic features. The second group of articles showed higher syntactic complexity, was engineering. The degree of syntactic complexity in medical sciences was lower than engineering and humanities. And the last group of RAs in the basic science field demonstrated the lowest degree of syntactic complexity. As indicated by the P value represented in Table 1, all of the linguistic variables showed to be fundamentally unique among the different orders with the exception of the primary measure of syntactic complexity; average sentence length. The majority of disciplines except humanities had the similar length of sentence.

The analysis of the moves revealed that M1, across the four corpora, was an obligatory Move. M1S1 (establishing the territory, claiming the importance of the research) and M1S3 (establishing a territory, reviewing items of the previous research) were observed to be available in both Persian and English local RAs, hence they were obligatory Moves. The same was true for international RAs. M1S1 and M1S3 occurred in more than 50% of international RAs. The results indicated non-significance of differences between M1S1 and M1S3 across four corpora. Both local and international researchers try to centralize their concern around a topic that they claim it is attractive. Across the four corpora, S2 (making topic generalization) was the only obligatory step in M2. This step was present in 40% of local RAs. While the frequency of occurrence of this step in international RAs was higher 60%. In both local and international RAs, the authors explicitly indicate a gap in previous studies done in the areas. It seems that international writers more willingly tried to locate the existing gap in order to justify their research. M2S2 (establishing a niche, indicating a gap) and M2S3 (establishing a niche, question raising) were the obligatory moves in local RAs. M2S2 (stating selected findings) as an obligatory step, was present in 100% of international and local RAs. M2S3 was additionally observed to be mandatory in both local and international corpora. In M3, S1 (outlining purposes and announcing present research) was the most frequent step among others. It was found in 67% of local RAs and 65% of international ones. It indicates that both local and international researchers prefer to outline their research purposes as opposed to expressing what the research is built up to exhibit. This move was considered as



mph.sdcongress.ir

هشتمین همایش بین المللی
مدیریت، روانشناسی و علوم انسانی
با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



obligatory in humanities and engineering English and Persian RAs. Generally, the results of this study as well as other findings referred to here are in accordance with Swales' (1990) observation that M3S1 is an obligatory step across RAs. Thus, it can be implied that all English and Persian research articles followed the same move pattern. Their writers picked Swales' s model as the criterion. The noticeable feature of almost all of the articles was the use of moves and steps of Swales' s model.

5. Discussion

Breaking down the conventional structure of RAs, cross-linguistically and cross-disciplinary, has gotten exceptional consideration by numerous researchers (Ahmad, 1997; Hirano, 2009). EFL analysts and the authors require more intensive and thorough information about the expository structure of RAs across various disciplines. Nonetheless, a couple of studies have been completed in the Iranian context compared to the other EFL contexts. Therefore, the current study tried to investigate the generic structure of review sections of research articles and linguistic devices which have the potential to distinguish among research articles in various disciplines. Linguistic properties of the review sections of research articles from four major disciplines which were indicators of text complexity, were conveyed to spread among these disciplines. Because of the very nature of the debates and the required reasonable examples went with disciplines and in accordance with cases made by Hyland (2000) and Widdowson (2007), it was hypothesized that the distinction in the essence of argumentation will show itself, somehow, at the simple surface level of writings identified with these disciplines. To do so, lexical complexity, syntactic complexity, and move complexity as three text complexity features in the review sections of research articles were centered upon to explore variation across four disciplinary territories; humanities, engineering, basic sciences, and medical sciences.

Lexical and syntactic complexity measures uncovered significant higher indices for humanities review sections. There was a significant difference between Persian and English articles in the field of humanities regarding lexical complexity and syntactic complexity. One main justification for failure in international humanities journal might be justified for this reason. The second most significant difference was found between engineering articles in Persian and their English counterparts. The Persian and English research articles demonstrated no significant difference in lexical and syntactic complexity. The observable element of the majority of the articles was the use of moves and steps of Swales' model.

The findings can be clarified in the light of the fundamental precepts proposed by Hyland (2000) and Widdowson (2007) who believed that language is a device for communicating thoughts and information which carries on distinctively in different social and linguistic contexts. Larger indices of complexity measures demonstrated by humanities articles in the study, were characteristic of the higher level of sophistication. According to previous research, more sophisticated writings show larger complexity indices. Hyland (2009) recommended that various disciplines have distinctive ways to present their approaches to display their arguments and ideas. He also asserted that successful authors have to be sophisticated writers in order to get the argument for their propositions. Verifying previous research, the linguistic metrics used in this study proved to have the potential to discriminate among various disciplines. Humanities discipline demonstrated the highest and basic sciences discipline showed the lowest degree of text complexity. Engineering and medical sciences revealed no remarkable variation regarding



mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



linguistic indices of their review sections. The overall proposed model of text complexity features applied a good level of contribution to discriminate disciplines.

The first contributing factor to differentiate disciplines was lexical density. The formula for calculating lexical density is the proportion of content words to the total number of words. The results revealed that texts written as research articles in humanities demonstrated higher lexical density indices. It is assumed that more lexically dense texts are also more sophisticated texts (Camiciottoli, 2007). Local and international authors who wrote for humanities articles produced the highest proportion of lexical or content words to the total number of words. This result was confirmed by the previous studies (Camiciottoli, 2007). The second contributing index was lexical diversity. It's a critical factor of complexity in grouping articles based on their relevant disciplines. Lexical diversity is a measure of vocabulary size. Various studies verified lexical diversity as a significant factor in language use (Laufer, 2003; Crossley, Salsbury, & McNamara, 2011). The findings indicated that argumentation in the humanities articles entails using more different kinds of words types.

The third factor was the ratio of subordinate structures. The ratio of subordination is a syntactic complexity index, which is computed by having the proportion of the subordinate structures to the combination of subordinate and coordinate structures. The previous study in this regard revealed the fact that as the complexity level of texts increases, they exhibit a higher proportion of the subordinate structures (Wolfe- Quintero et al., 1998; Iwashita, 2005). The findings of this study also proved a higher ratio of subordination for humanities articles' review sections. The fourth factor was the average sentence length. This factor measures the average number of words per sentence and the number of different kinds of sentences, including; simple, compound, complex, and complex-compound sentences. The authors of humanities' sub-fields tended to use lengthier sentences which were mainly simple and complex ones.

The last contributing index to distinguish among various disciplines was move complexity. In this study, Swales's CARS model was used and no remarkable difference was demonstrated between English and Persian research articles. M1 was recognized as an obligatory move. M1S1 and M1S3 were found to be present in local RAs, the same was true for international RAs. Swales (1990) states that the value function of M1S1 is to persuade the members of the discourse community to accept that the research which is to be presented is taken from a lively, significant and well-established research area. Peters (1997) reported that the main function of M1S3 is to provide justification for the research being conducted. Similarly, Hart (2001) noticed that this step helps the students to determine the scope of their research and enables them to locate the existing gaps in the previous studies. Samraj (2002) argues that M1S3 depicts the importance of the study being reported. S2 was the only obligatory step in M2. In both English and Persian articles, authors explicitly indicate a gap in previous studies done in the areas they were investigating; however, it seems that international writers more willingly tried to locate the existing gap in order to justify their research. These findings are in line with Kwan (2006) and Samraj (2005) who, in their studies, found that indicating a gap or what they called "the dearth" or "dearth of relevant studies" is present in RAs and doctoral theses. In M3, S1 was the most frequent step among others. It indicates that both local and international researchers prefer to outline their research purposes instead of stating what the research is established to present. Generally, the results of this study are in line with Swales' (1990) observation that M3S1 is an obligatory step of RAs. In conclusion, the general steps used in the review sections of local and international articles were "Claiming centrality (or) setting the ground by term definition, elaboration, or exemplifying, reviewing



mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



related research, Indicating a gap, and Outlining purposes". As the final remark, it can be concluded that rhetoric structure of RAs seems to be universal. Both local and international authors follow the same structure. These results support the Widdowson's (1979) beliefs in the universality of rhetorical structure of RAs.

6. Conclusion

This study aimed to investigate the linguistic measures which could be applied in discriminating various disciplines based on the reviews written in four major fields of study. In line with previous research, lexical density, lexical diversity, average sentence length, and the ratio of subordination turned out to be significantly different in various disciplines. These linguistic indices were significantly higher in the humanities articles; meanwhile, the basic and medical science groups indicated no significant difference with respect to the linguistic measures. The results showed that the strongest factor in disciplines was lexical density followed by a ratio of subordination. The demonstrated differences can be attributed to the very nature of the disciplines. These findings imply that EAP and ESP teachers need to take into account the major fields of study and the potential writing conventions of those disciplines.

References

- [1] Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26, 390-395.
- [2] Bhatia, V. K. (1993). *Analyzing genre: Language use in a professional setting*. London and New York: London Group.
- [3] Bhatia, V. K. (1999). Integrating products, processes, purposes, and participants in professional writing. In C. N. Candlin & K. Hyland (Eds.), *Writing: Texts, processes, and practices* (pp. 21-39). London: Longman.
- [4] Bhatia, V. K. (2001). *Analyzing genre: Some conceptual issues*. In M. Hewings (Ed.), *Academic writing context* (pp. 79-92). Birmingham: University of Birmingham.
- [5] Brett, P. (1994). A genre analysis of the results section of sociology articles. *English for Specific Purposes*, 13(1), 47-59.
- [6] Bachman, L. F. (1990). *Fundamental Consideration in Language Testing*. Oxford: Oxford University Press.
- [7] Bhatia, V. K. (2001). *Analyzing genre: Some conceptual issues*. In M. Hewings (Ed.), *Academic writing context* (pp. 79-92). Birmingham: University of Birmingham.
- [8] Boote, D. N., and Beile, P. (2005), "Scholars before researchers: On the centrality of the dissertation Literature review in research preparation", *Educational Researcher*, Vol. 34 No. 6, pp. 3-15.
- [9] Byrnes, H. (2009). Emergent L2 German writing ability in a curricular context: A longitudinal study of grammatical metaphor. *Linguistics and Education*, 20, 50-66. <http://dx.doi.org/10.1016/j.linged.2009.01.005>.
- [10] Chen, M., & Zechner, K. (2011). Computing and evaluating syntactic complexity features for automated scoring of spontaneous non-native speech. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics* (pp.722-731).
- [11] Chun, C., & Pennebaker, J. (2007). The psychological functions of function words. In: K. Fiedler (Ed), *Social Communication*. New York, NY: Psychology Press.



mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



- [12] Crawford Camiciottoli, B. C. (2004). Interactive discourse structuring in L2 guest lectures: Some insights from a comparative corpus-based study. *Journal of English for Academic Purposes* 39- 54.
- [13] Crawford Camiciottoli, B. C. (2007). *The Language of Business Studies*. Amsterdam: John Benjamin Publishing Company.
- [14] Crossley, S.A., &McNamara, D. S. (2009). Computationally assessing lexical differences in second language writing. *Journal of Second Language Writing*,18(2), 119-135.
- [15] Crossley, S. A., Salsbury, T., & McNamara, D.S. (2011). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing*, 28,561-580.
- [16] Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26, 59–84.
- [17] Ellis, R. (2005). Planning and task-based performance: Theory and research. In R. Ellis (Ed.), *Planning and task performance in second language* (pp. 3-34). Amsterdam: John Benjamins.
- [18] Engber, C. A. (1995). The relation of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*,4(2),139-155.
- [19] Fallahi, M. M., & Erzi, M. (2003). Genre analysis in language teaching: An investigation of the structure of the discussion section of language teaching-journal articles. *Iranian Journal of Applied Linguistics*, 6(1), 69-81.
- [20] Foster, P. &Tavakoli, P. 2009. Native speakers and task performance: Comparing effects on complexity, fluency and lexical diversity. *Language Learning* 59(4): 866–896.
- [21] Hart, C. (1988). *Doing a literature review: Releasing the social science research imagination*. Thousand Oaks: Sage
- [22] Hart, C. (2001). *Doing a literature search: A comprehensive guide for the social sciences*. London: Sage Publications.
- [23] Hirano, Eliana. "Research Article Introduction in English for Specific Purposes: A Comparison between Brazilian Portuguese and English." *English for Specific Purposes* 28 (2009): 240-50.
- [24] Hunt, K. (1965). *Grammatical structures written at three grade levels*. NCTE Research Report No.3. Champaign, IL: NCTE.
- [25] Hyland, K. (1996). Talking to the Academy: Forms of hedging in science research articles. *Written Communication*, 13(2), 251-281.
- [26] Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. London, UK: Longman.
- [27] Hyland, K. (2003). Self-citation and self-reference: Credibility and promotion in an academic publication. *Journal of the American Society for Information Science and Technology*, 54(3), 251-259.
- [28] Hyland, K. (2008). *As can be seen: lexical bundles and disciplinary variation*. Center of Academic and Professional Literacies, Institute of Education, University of London, 20 Bedford.
- [29] Hyland, K. (2009). *Academic Discourse* London: Continuum.
- [30] Iwashita, N. (2005). An investigation of lexical profiles in performance on EAP speaking tasks. *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 3,101-112.
- [31] Jarvis, S. (2013). Defining and measuring lexical diversity. In S. Jarvis, & M. Daller (Eds.), *Vocabulary knowledge: Human ratings and automated measures* (pp. 13–45). Amsterdam: John Benjamins.
- [32] Kwan, B. (2006). The schematic structure of literature reviews in doctoral theses of applied linguistics. *English for Specific Purposes*, 25, 30-55.



mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



- [33] Koda, K. (1993). Task-induced variability in FL composition: Language-specific perspectives. *Foreign Language Annals*, 26, 332-346.
- [34] Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing*, 17, 48-60.
- [35] La Brant, L. L. (1934). A study of Certain Language Developments in Children, *Genetic Psychology*, 14,387-491
- [36] Laufer, B., & Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- [37] Laufer, B. (2003). The influence of L2 on L1 collocational knowledge and on L1 lexical diversity in free written expression. In V. Cook (Ed), *Effects of Second Language on the First* (pp.19-31) Clevedon, UK: Multilingual Matters.
- [38] Lu, Xiaofei (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4): 474-496.
- [39] Lu, X (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly* 45(1): 36–62.
- [40] Li, L., & Ge, G. (2009). Genre analysis: Structural and linguistic evolution of the English-medium medical RA (1985–2004). *English for Specific Purposes*, 28, 93-104.
- [41] Laufer, B., & Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- [42] Laufer, B. (2003). The influence of L2 on L1 collocational knowledge and on L1 lexical diversity in free written expression. In V. Cook (Ed), *Effects of Second Language on the First* (pp.19-31). Clevedon, UK: Multilingual Matters.
- [43] Malvern, D. D., Richards, B. J., Chipere, N., & Duran, P. (2004). Lexical diversity and language development: Quantification and assessment. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.10.1057/9780230511804.
- [44] Malvern, D.D., Richards, B.J., Chipere, N., & Durán, P. (2004) . Lexical diversity and language development. New York: Palgrave MacMillan
- Martin-Martin, P. (2003). A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purpose*, 22(1), 25-43.
- [45] McCarty, P. M. (2005). An Assessment of the Range and Usefulness of Lexical Diversity Measures and the Potential of the Measure of Textual, Lexical Diversity (MTLD). Unpublished Ph.D. dissertation, The University of Memphis.
- [46] Melander, B., Swales, J. M., & Fredrickson, K. M. (1997). Journal abstracts from three academic fields in the United States and Sweden: National or disciplinary proclivities? In A. Duszak (Ed.), *Culture and style of academic discourse* (pp. 251-272). New York, NY: Mouton de Gruyter.
- [47] Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492–518.
- [48] Ong, J., & Zhang, L., J. (2010). Effects of task complexity on fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing*, 19, 218-233.
- [49] Ozturk, I. (2007). The textual organization of research article introduction in applied linguistics: Variability within a single discipline. *English for Specific Purposes*, 26(1), 25-38.
- [50] Peters, R. L. (1997). *Getting what you came for* (revised.). New York: Noonday.
- [51] Pho, P. D. (2008). Research article abstracts in applied linguistics and educational technology: A study of linguistic realizations of rhetorical structure and authorial stance. *Discourse Studies*, 10(2), 231-250.



mph.sdcongress.ir

هشتمین همایش بین المللی مدیریت، روانشناسی و علوم انسانی با رویکرد توسعه پایدار

8th International Conference on
Management, Psychology & Humanities with sustainable development approach



- [52] Polio, C. (2001). Research methodology in second language writing research: The case of text-based studies. In T. Silva & P. K. Matsuda (Eds.), *On second language writing* (pp. 91-115). Mahwah, NJ: Lawrence Erlbaum.
- [53] Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics*, 45, 193-213.
- [54] Samraj, B. (2002). Introductions in research articles: Variations across disciplines. *English for Specific Purposes*, 21(1), 1-17.
- [55] Samraj, B. (2005). An exploration of a genre set: Research article abstracts and introduction in two disciplines. *English for Specific Purposes*, 24(2), 141-156.
- [56] Swales, J. M. (1981). *Aspects of article introductions*. Birmingham: The University of Aston.
- [57] Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- [58] Swales, J. M. (2004) *Research Genres: Explorations and applications*, Cambridge: Cambridge University Press.
- [59] Swales, J. M., & Feak, C. B. (2009). *Abstracts and the writing of abstracts*. Ann Arbor, MI: University of Michigan Press.
- [60] Swales, J. M. (2009). When there is no perfect text: Approaches to the EAP practitioner's dilemma. *Journal of English for Academic Purposes*, 8, 5-13.
- [61] Scott, M. (2009). *Wordsmith Tools Liverpool: Lexical Analysis Software*.
- [62] Shaw, P., & Liu, E. (1998). What develops in the development of second-language writing? *Applied Linguistics*, 19, 225-254.
- [63] Vyatkina, N. (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *The Modern Language Journal*, 96, 576-598. <http://dx.doi.org/10.1111/j.1540-4781.2012.01401.x>.
- [64] Vyatkina, N. (2013). Specific syntactic complexity: Developmental profiling of individuals based on an annotated learner corpus. *The Modern Language Journal*, 97, 11-30. <http://dx.doi.org/10.1111/j.1540-4781.2012.01421.x>.
- [65] Widdowson, H. G. (1979). *Explorations in Applied Linguistics*. Oxford: Oxford University Press.
- [66] Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity (Report No.17)*. Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- [67] Yu, G. (2010). Lexical diversity in writing and speaking task performances. *Applied Linguistics*, 31, 236-259.