

Lexical Complexity and Discourse Markers in Soft and Hard Science Articles

Javad Gholami, Zahra Mosalli and Shiva Bidel Nikou

Urmia University, Iran

Abstract: Discourse markers (DMs) are linguistic elements that index different relations and coherence between units of talk. Most research on the development of these forms has focused on conversations rather than written materials. This study investigated the development of DMs in the abstracts of ESP articles along with their lexical complexity. To this end, research papers in psychology and tourism as soft sciences and physics and engineering as hard sciences were analyzed. Thirty articles were selected randomly for each area. The frequency of discourse markers was examined and their rate of lexical complexity was estimated based on the percentage of lexical words to total words in whole texts and also based on the proportion of lexical to function words. The results of the analyses revealed that *contrastive* and *additive* markers were the most frequently used ones in these article abstracts. As far as lexical density was considered, there was no outstanding difference between these fields of hard and soft sciences. In the case of the proportion of lexical to function words, articles on psychology were less complex than the ones in the other fields.

Key words: English for Academic Purposes (EAP) • Discourse markers • Lexical complexity • Soft sciences
• Hard sciences

INTRODUCTION

The need to communicate in specialized contexts or domains, such as academic and scientific disciplines at university, is greatly emphasized in current academic research areas. The emphasis is often placed on an effective linguistic development for research purposes [1]. The new venues and scopes suggested for ESP (English for Specific Purposes) point to the importance of conveniently valuing and assessing the development of specialized languages, special linguistic features and lexical complexity in different areas. According to Fuertes-Olivera [2], the other newly established scope in ESP course has included the significance of suitably assessing the development of lexical focus. The language which is used in academic writing such as research writing is of importance for academic people such as students, professors and research writers. Thus, the language and its characteristics (like lexical complexity and discourse markers) can be a good indicator of successful communication in ESP/EAP settings. One of the reasons for the emergence of ESP is revolution in linguistics. Whereas traditional linguists set out to describe the

features of language, revolutionary pioneers in linguistics began to focus on the ways in which language is used in real communication.

Communication in written works can be appropriately established through the use of discourse markers. Discourse markers can be good indicators of cohesion and coherence in written outputs. Cohesion and coherence are two factors in the English text. In order to successfully communicate in academic writing, it is necessary for students to learn about cohesive and coherent devices. Halliday and Hasan [3] discuss how reference, substitution, ellipsis, conjunction and lexical cohesion create cohesion in the text. They indicate that reference, substitution and ellipsis are clearly grammatical; lexical cohesion is, as the name implies, lexical; and conjunction is on the borderline of the grammatical and the lexical. The conjunction they mentioned refers to not only conjunctions but also adverbs and prepositional phrases with conjunctive functions.

Text and discourse comprehension is central to a vast range of human enterprises. For text information to effectively guide people's behavior, it is necessary that the reader scrutinize the accuracy of text ideas. There is

extensive evidence that readers are sensitive to the congruence between the current portion of a message and its antecedents [4]. Readers can detect discrepancies on the basis of relatively apparent mismatches at the surface level of discourse [5]. However, they also notice inference-based inconsistencies stemming from numerous dimensions of the text situation model [6]; including logical [7] and causal [8] facets of text.

Zare and Noordin [9] believe that reading plays a crucial role in English language teaching and learning. Reading is one of the most important goals of many EFL/ESL learners and the primary purpose for reading is comprehension. In this regard Anderson and Lynch [10] state that reading comprehension depends on three main sources of knowledge which are schematic knowledge, contextual knowledge and systematic knowledge and knowledge about the discourse markers is considered as a systematic knowledge. Furthermore, Widdowson [11] believes that an influential factor in comprehension of a text is the structure of it and two properties of any text are cohesion and coherence [12]. Coherence is an underlying semantic relation which turns the words, sentences, or propositions into a unified understandable whole and is achieved by interpreting each individual sentence and relating these interpretations to one another [13]. But cohesion as 'surface-level ties' links separate phrases, clauses, sentences and even paragraphs into a unified discourse [14]. Additionally, Nunan [15] emphasizing the term cohesive devices, believes that "coherent texts are distinguished from random sentences by the existence of certain text-forming, cohesive devices" (p. 59). These words are described as "cohesive devices," as they create links across the boundaries of mere fragments, or can chain related items together. In a nutshell, text structure is assumed to be a decisive factor in revealing the writers' attitude while text structure itself is responsible to certain meta-discourse features. To illuminate the point, meta-discourse features are those facets of a text which make the organization of the text explicit, provide information about the writer's attitude toward the text content and engage the reader in the interaction [16].

Another area in academic research is the study on lexical characteristics of ESP articles. They play important roles in identifying different lexical diversities in hard and soft science articles. Studies on lexical items in academic research have different subareas. Frequency (high, mid and low) of function or content words, accuracy of them, predictability of lexical items production and broadly speaking, phonological, ethological and lexical processes are the major investigation lines and areas of enquiry of many academic studies.

According to Diaz and Mc Carthy [17], content and function words are two classes of words that follow the same rules of English orthography and phonology, but differ markedly in their role in language and in the degree to which they represent meaning. Content words are nouns, verbs, or adjectives that convey semantic information. They are often, but not always, associated with a physical object (e.g., house, table), have many associations to other words (e.g., dog, cat) and have meaning independent of context. The number of content words in a lexicon can always be increased as new objects or concepts are created or discovered. Thus, content words are often referred to as open-class words. In contrast, function words have linking and syntactic functions in context (e.g., although, while), but have few inherent associations to other words when presented outside of a sentence. The number of function words is generally fixed within a lexicon and so they are also termed closed-class words. While content words retain their semantic properties independent of context, the degree to which function words engage syntactic processes outside of a sentential context is debated. Brill [18] made a distinction between content and function words. He did this by using parts of speech to label the words and then editing the resulting word labels by hand. Some changes appeared in his work, for example, *okay* and *well* changed from modifier to discourse markers. Diaz and Mc Carthy [17] examined the words likely occur in more than one category such as *so* and verb forms like *have* and then corrected their labels. They classified the words into two categories content versus function words. In their classifications, the function-word category included discourse markers, conjunctions, existential there, pronouns and other prepositions, articles, quantifiers, demonstratives, verb auxiliaries and verb particles. The remainder, nouns, verbs, adverbs and adjectives, made up the content-word category.

In this study we attempted to investigate the other rarely touched research area on lexical items, that is lexical complexity. In assessing the lexical complexity, we utilized Diaz and Mc McCarthy's [17] classifications of function and content words. Lexical complexity is classified into lexical density and the proportion of lexical words to function words. Lexical density which is also called type-token ratio is a measure of the relationship between the number of words with lexical as opposed to grammatical properties as a percentage of the total number of words in a text. Lexical density is sometimes used as a measure of difficulty of a passage or text. It is the ratio of lexical words to the total number of words in a text multiplied by 100.

Literature Review: There are many research works about the cohesive devices in different academic fields. For example, Li [19] in a study pointed out that conjunction is an indicator of the quality of fluency and complexity of learners' oral production and it was shown that learners had less ability to produce conjunction in their speaking. Moreover, in another study by Spivey [20], the findings revealed that writers utilized a meaning-making process that involved organizing ideas as they read and wrote, selecting from readings and integrating ideas based on inferences from texts. Writers used cohesion markers to make connections within and across texts. These findings led Spivey [20] to conclude that organization, selection and connection "are the very basis of reading, writing and learning in almost any domain knowledge" (p. 191).

Different studies have made different and often unique, selections of text features. Nevertheless, though made operational in various ways, a number of broad aspects tend to recur across these studies: communicative effectiveness, content, organization, language use and fluency/length. Moreover as far as the study on vocabulary items was taken into account, the areas of study tended to be narrower such as the frequency, accuracy and range of lexical items. As all writers were composing essays in English, not their first language, the issue of language choice in the think-aloud sessions required consideration [21]. In addition, the language choice differs from one major of field to another. For instance, in many natural science and engineering disciplines, empirical or quantitative data are incorporated into reasons that are used to support the authors' claims. When students write up empirical studies, they commonly produce essays based on some kind of theory and empirical data. In the social sciences, the author must typically argue for the theory or framework as well as conclusions (i.e., claims) based on data. A particular genre is the laboratory report with highly specified sections (abstract, introduction, methods, results, discussion). It is fruitful to think of lab reports as entailing a set of prescribed yet tacit arguments. The introduction supports the unstated claim "this is a worthwhile question," and the methods section supports the unstated claim "these methods are valid." The results and discussion sections together make an argument with the empirical results providing the reasons and the discussion setting forth key claims. In another vein, observation papers require the writer to make empirical observations, which are then used to support arguments in the paper [22].

A great deal of research on discourse markers has been carried out during the past two decades [23-26]. Although other terms such as "discourse particles,"

"connectives," "pragmatic expressions" or "pragmatic markers" are preferred by some researchers, the term "discourse markers" (DMs) is more commonly employed by researchers who work on English discourse [27]. Therefore, the term "discourse markers" will be used in this article.

Waring [28] demonstrated the intricate operation of the adverb *also* in actual interaction at a level of detail that dictionary definitions failed to be captured. Using primarily a conversation analytic framework in examining two data corpora, which included a series of graduate seminar discussions and television roundtable discussions, he argued that the semantic features of *also* are strategically deployed to accomplish complex interactional goals in a disjunctive or disaffiliative environment. In a disjunctive environment, *also* can be invoked to legitimize one's speaking rights to get the floor. In a disaffiliative environment, *also* can be mobilized to either soften or strengthen a disaffiliative action in subsequent talk. These practices of *also* are accounted for in part by the tensions between coherence and continuation [29, 30].

In an attempt to explore the stances, intentions and ideologies of the United States and China that underlie news stories of the spy-plane event that occurred in April 2001, Cheng [31] investigated 94 news stories from a US online medium (CNN.com) and 15 from a Chinese online medium (ChinaOnline.com) during the 12-day period between the air collision and the return of the 24 US crew members (1-12 April 2001). By interpreting utterances containing the *if*-clause + the situational context, the implicature of 'uncertainty' was revealed through one of the key pragmatic implicatures that was linked to the conditional marker 'if'.

Dailey-O'Cain and Liebscher [32] focused on learners' use of discourse markers in one advanced German language classroom. In analyzing the data for functional distribution of pairs of German and English discourse markers as used by the students, they found that the discourse markers *so* and *also* had specialized functionally on the level of the entire classroom community of practice.

Bell [33], using an 8 million-word corpus together with introspected examples, examined the concessive cancellative discourse markers *nevertheless*, *still* and *yet*. The corpus yielded 254 tokens of *nevertheless/nonetheless*, 802 tokens of discourse marker *yet* and 262 tokens of discourse marker *still*. Those markers were distinguished according to three interconnected criteria: variability of scope, speaker perspective and degree of concession. With regard to the

degree of concession, *nevertheless*, *still* and *yet* were again considered in terms of agreement with *nevertheless* signaling the strongest degree of concession and *yet* the weakest. It was no coincidence that the clines of scope and concession were inversely related: the vaguer the instruction carried by a concessive marker, the greater its ability to operate globally and conversely, the more detailed the instruction, the less its ability to operate globally.

Bolden [34], using the methodology of conversation analysis, examined a large corpus of recorded conversations to explicate the role of *so* in implementing incipient actions. The analysis focused on the use of *so* for prefacing sequence-initiating actions (such as questions) and demonstrated that speakers deployed this preface to indicate the status of the upcoming action as 'emerging from incipiency' rather than being contingent on the immediately preceding talk. *So* prefacing was recurrently used in contexts where the activity being launched was relevantly pending. Additionally, speakers could use *so* to characterize and constitute a particular action as advancing their interactional agenda. The article showed that this marker is a resource for establishing discourse coherence and, more fundamentally, for accomplishing understanding.

Verdonik and Zgank [35] used the results of an analysis of discourse marker use in two different conversational genres (telephone conversation and television interviews) in an attempt to examine the impact of context on the use of discourse markers, generalized for each analyzed genre. In the first stage of the analysis, they observed important differences between the two genres: discourse markers were far more frequently used in telephone conversations than in television interviews. In the second stage of the analysis, they identified several contextual factors which contributed to the differences in the use of discourse markers.

Chen and Ge [36] conducted a lexical study on the word frequency and the text coverage of the 570 word families from Coxhead's Academic Word List (AWL) in medical research articles (RAs) based on a corpus of 50 medical RAs written in English with 190425 running words. By computer analysis, they found that the text coverage of the AWL words accounted for around 10.07% in English medical RAs, that 292 (51.2%) out of the 570 AWL word families were frequently used in English medical RAs and that the academic words used in English medical RAs distributed dispersedly throughout a whole RA, accounting on the average for around 10% text coverage with slight difference among the five sections of

a medical RA (Abstract, Introduction, Materials and Methods, Results and Discussion). From these findings, they concluded that: (a) academic vocabulary, with a high text coverage and dispersion throughout a medical RA, is an important set of word items in medical RAs; (b) the AWL, a list of academic vocabulary representing academic word families across a wide range of subject disciplines, is far from complete in representing the academic words frequently used in medical RAs; and (c) the different coverage of academic words in the different sections in a medical RA, together with the role each section is supposed to play in a medical RA, indicates that academic words to a great extent serve some rhetorical functions in academic texts, as seen in the medical RAs in their study.

In another study by Fuertes-Olivera [37], the lexical gender in written business English was taken into account. His study investigated lexical gender in specialized communication. The key method of analysis was that of forms of address, professional titles and 'generic man' in a 10 million word corpus of written Business English. Mixed results were found. On the one hand, the 'male-as-norm' principle contributes to reinforcing typical gender stereotypes: for example, for each woman referred to in the corpus, there were more than 100 occurrences for man. On the other hand, advocates of non-sexist English had also influenced written Business English: for example, *Ms* was more than 9 times as frequent as *Mrs.* and *Miss*, which sustained the claim that equates *Ms* with professional settings.

The present study attempted to consider and investigate the lexical complexity and discourse markers in different academic fields in order to see which category (hard science or soft science) is more lexically complex. The number of discourse markers and most frequent type of them is the other major line of investigation in our study.

MATERIALS AND METHODS

Four academic fields, psychology and tourism as soft sciences and engineering and physics as hard sciences were chosen to be studied in the present work. One-hundred twenty articles, that is, 30 articles in each field were investigated. The abstracts of the articles were analyzed to find out their rate of lexical complexity and the frequency of discourse markers used in them. Lexical complexity was measured through the lexical density and the proportion of lexical (content) words to function (grammatical) words. Lexical density was calculated through the following formula:

Table 1: Lexical Complexity (lexical density and proportion of lexical words to function words)

Area		Number of articles	Number of words	LD %	L/F
Soft science	Psychology	30	4249	59.20	0.93
	Tourism	30	4058	56.77	1.55
Hard science	Physics	30	4209	58.00	1.51
	Engineering	30	4355	57.32	1.64
Total	-	120	16871	57.89	5.63

Table 2: Discourse Markers in soft and hard sciences

Types of DMs	Soft science		Hard science		Total
	Psychology	Tourism	Physics	Engineering	
Contrastive	18	8	15	7	48
Additive	11	10	15	10	46
Enumerative	10	10	3	-	23
Logical sequence	7	6	3	6	22
Explicative	-	2	-	2	4
Illustrative	2	-	-	-	2
Total markers	48	36	36	25	145

$$lexical\ density = \frac{lexical\ words}{total\ words} \times 100$$

Function words (or grammatical words or semantic words or structure-class words) are words that have little lexical meaning or have ambiguous meaning, but instead serve to express grammatical relationships with other words within a sentence, or specify the attitude or mood of the speaker. They signal the structural relationships that words have to one another and are the glue that holds sentences together. Thus, they serve as important elements to the structures of sentences.

Words that are not function words are called content words (or open class words or lexical words or auto semantic words): these include nouns, verbs, adjectives and most adverbs, although some adverbs are function words (e.g., *then* and *why*). Dictionaries define the specific meanings of content words, but can only describe the general usages of function words. By contrast, grammars describe the use of function words in detail, but treat lexical words in general terms only.

Data Analysis: The results of the investigation of lexical complexity of the abstracts are shown in Table 1.

Based on Mackay [38], the types and number of DMs used in the articles are presented in Table 2.

DISCUSSION AND CONCLUSION

By analyzing the extent to which lexical complexity is present in the abstracts of hard science (engineering and physics) and soft science (tourism and psychology)

articles, we have drawn two main conclusions regarding two factors of lexical complexity, that is the proportion of content (lexical) words to function (grammatical) words and the lexical density. First, as clearly illustrated in Table 1, the proportion of lexical words to function words, as a factor of lexical complexity, was different for different fields. However, this difference was not outstanding, especially among tourism, physics and engineering articles. Abstracts of articles on psychology (0.93) were somewhat different from other fields' article abstracts. In this regard, the highest amount was for engineering (1.64). Secondly, the lexical density, as another factor of lexical complexity, was the highest in psychology and the least in tourism, indicating the most difficult texts for psychology and the easiest ones for tourism. As far as the lexical complexity was considered, the results of present study were based on the methods used in a study by Kuiken and Vedder [39]. In that study, the lexical variation in the texts based on the complex task as measured by the type-token ratio (TTR1) was significantly larger than in those based on the non-complex task; however, this finding was not confirmed by the alternative type-token ratio (TTR2).

As far as discourse markers are considered, the total number of them in the articles on psychology was 48 (the highest) and on engineering 15 (the least). As a whole, it should be stated that *contrastive markers*, mainly *however* and *additive markers*, mainly *also*, were the most frequent markers used in these abstracts. The frequency of *however* can be, in a way, because of the frequency of contrastive genre among these articles. According to Waring [28], *also* has the following two

meanings: (1) in addition; too; besides (e.g. 'He's young and he's also brilliant.');

and (2) likewise; in the same manner (e.g. 'Since you're having another cup of coffee, I'll have one also.').

The first meaning 'in addition', 'too' or 'besides' ('additive' henceforth) involves some sort of listing and the features being enumerated are bound by a common topic. In 'He's young and he's also brilliant', for example, one is counting the qualities of a man - the common topic. The second meaning 'likewise' underscores similarities of some sort. In 'Since you are having another cup of coffee, I'll have one also', is identical act of 'having another cup of coffee' that warrants the use of *also*. Waring [28] also believes that in a disjunctive environment, *also* can be invoked to legitimize one's speaking rights to get the floor. In a disaffiliative environment, *also* can be mobilized to either soften or strengthen a disaffiliative action in subsequent talk. These practices of *also* are accounted for in part by the tensions between coherence and continuation. Dailey-O'Cain and Liebscher [32] in their study found that discourse markers *so* and *also* had specialized functionally on the level of the entire classroom community of practice. It should be added that *so* belongs to *logical sequence* category of DMs which were the next frequent markers after *contrastive* and *additive markers* in these texts. This result is somehow in line with the results of the study done by Dailey-O'Cain and Liebscher [32]. However, *explicative* and *illustrative* markers were the least frequent ones in these abstracts which can be because of the conciseness of abstracts and the necessity for being to the point.

The work described here only focused on larger and fewer DMs, whereas other types of manipulation and other counts of DMs can be regarded as well. Further research can investigate whether the results of this study are the same about spoken contexts, for example in seminars, to see if the same results come out in two different contexts of written and spoken form. In addition, one can investigate other sections of the articles as well as the whole text. Besides, various text genres, their uses of DMs and their lexical complexity can also give a broader perspective of the findings of this study. Moreover, further research can investigate the other factors of lexical complexity such as the proportion of technical terms (vocabulary items) to non-technical words or the percentage of low-frequency words to high frequency ones. In summary, the frequency, accuracy and complexity of lexical words can be the major line of investigation of further research in future. These areas of research can be done in different fields of study in

academic levels and also in different parts of theses and research articles (such as abstract, introduction, methodology, etc.).

The results of this study demonstrate types and numbers of discourse markers and also lexical complexity in different academic fields. Thus, they can be used in writing assessments in those areas.

REFERENCES

1. Bricall Report, 2000. Informe 'Universidad 2000'. El País, pp: 14.
2. Fuertes-Olivera, P.A., 2007. A corpus-based view of lexical gender in written business English. *English for Specific Purposes*, 26: 219-234.
3. Halliday, M.A.K. and R. Hasan, 1976. *Cohesion in English*. London: Longman.
4. Ferretti, T.R., M. Singer and C. Patterson, 2008. Electrophysiological evidence for the time-course of verifying text ideas. *Cognition*, 108: 881-888.
5. Long, D.L. and J.L. Chong, 2001. Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27: 1424-1429.
6. Zwaan, R.A., J.P. Magliano and A.C. Graesser, 1995. Dimensions of situation-model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21: 386-397.
7. Lea, R.B. and E.J. Mulligan, 2002. The effect of negation on deductive inferences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28: 303-317.
8. Suh, S. and T. Trabasso, 1993. Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols and recognition priming. *J. Memory and Language*, 32: 279-300.
9. Zare, P. and N. Noordin, 2001. The relationship between language learning strategy use and reading comprehension achievement among Iranian undergraduate EFL learners. *World Applied Sciences Journal*, 13(8): 1870-1877.
10. Aqsha Lubis, M., M. Mohamad, S. Hussin, Z. Amir, R. Din and Hamdan, 2011. Learners' perceptions toward group discussion via face-to-face and online forum in reading comprehension and learning ethic. *World Applied Sciences Journal*, 15: 23-28.
11. Widdowson, H.G., 1978. *Teaching language as communication*. Oxford: Oxford University Press.

12. Jafarinejad, R. and M. Tavakoli, 2011. Investigating the relationship between discourse markers, language proficiency and reading comprehension: A case of some Iranian university students. *Procedia Social and Behavioral Sci.*, 15: 1526-1530.
13. Van Dijk, T.A., 1977. *Text and Context: Explorations in the Semantics and Pragmatics of Discourses*. New York: Longman.
14. Gumperz, J.J., H. Kaltman and M.C. O'Connor, 1984. Cohesion in spoken and written discourse: ethnic style and the transition to literary. In D. Tannen, (ed.), *Coherence in spoken and written discourse*. Norwood, NJ: Ablex, pp: 3-19.
15. Nunan, D., 1993. *Introducing Discourse Analysis*. London: Penguin Group.
16. Intaraprawat, P. and M.S. Steffenson, 1995. The use of metadiscourse in good and poor ESL essays. *J. Second Language Writing*, 4(3): 253-273.
17. Diaz, M.T. and G. McCarthy, 2009. A comparison of brain activity evoked by single content and function words: An fMRI investigation implicit word processing. *Brain Res.*, 1982: 38-49.
18. Brill, F., 1995. Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging. *Computational Linguistics*, 21: 543-566.
19. Li, M., 2008. University English learners' speaking competence in their monologic production. *Polyglossia*, pp: 15.
20. Spivey, N., 1997. *The constructivist metaphor: Reading, writing and making of meaning*. San Diego, CA: Academic Press.
21. Machón, R.M., L. Murphy and J. Roca De Larios, 2005. Using concurrent protocols to explore L2 writing processes: Methodological issues in the collection and analysis of data. In P.K. Matsuda and T. Silva, (Eds.), *Second language writing research* Mahwah, NJ: Lawrence Erlbaum, pp: 191-205.
22. Wolfe, C.R., 2011. Argumentation across the curriculum. *Written Communication*, 28(2): 193-219.
23. Aijmer, K., 2002. *English discourse particles. Evidence from a corpus*. Amsterdam/Philadelphia: John Benjamins.
24. Fraser, B., 1990. An approach to discourse markers. *J. Pragmatics*, 14: 383-395.
25. Schiffrin, D., 1987. *Discourse Markers*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
26. Schourop, L., 1985. *Common Discourse Particles in English Conversation*. Garland, New York.
27. Liao, S., 2009. Variation in the use of discourse markers by Chinese teaching assistants in the US. *J. Pragmatics*, 41: 1313-1328.
28. Waring, H.Z., 2003. 'Also' as a discourse marker: its use in disjunctive and disaffiliative environments. *Discourse Studies*, 5(3): 415-436.
29. Linell, P., 1998. *Approaching dialogue. Talk and interaction in dialogical perspectives*. Amsterdam: John Benjamins.
30. Tracy, K. and J.P. Moran, 1983. *Conversational Relevance in Multiple-Goal Settings*. In R.T. Craig and K. Tracy, (Eds.), *Conversational Coherence*, Beverly Hills, CA: Sage, pp: 116-35.
31. Cheng, M., 2002. The Standoff - what is unsaid? A pragmatic analysis of the conditional marker 'if'. *Discourse and Society*, 13(3): 309-317.
32. Dailey-O'Cain, J. and G. Liebscher, 2006. Language learners' use of discourse markers as evidence for a mixed code. *International J. Bilingualism*, 10(1): 89-109.
33. Bell, D.M., 2010. Nevertheless, still and yet: Concessive cancellative discourse markers. *J. Pragmatics*, 42: 1912-1927.
34. Bolden, G.B., 2009. Implementing incipient actions: The discourse marker 'so' in English conversation. *J. Pragmatics*, 41: 974-998.
35. Verdonik, D. and A. Zgank, 2008. The impact of context on discourse marker use in two conversational genres. *Discourse Studies*, 10(6): 759-775.
36. Chen, Q. and G.C. Ge, 2007. A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs). *English for Specific Purposes*, 26(4): 502-514.
37. Fuertes-Olivera, P.A., 2007. A corpus-based view of lexical gender in written business English. *English for Specific Purposes*, 26: 219-234.
38. McKay, R., 1987. Teaching the information gathering skills. In Long and Richards 1987.
39. Kuiken, F. and I. Vedder, 2008. Cognitive task complexity and written output in Italian and French as foreign languages. *J. Second Language Writing*, 17: 48-60.