A Computerized Corpus Analysis of the Use of Pragmatic Markers in Native and Non-Native Corpora

Ali Şükrü ÖZBAY
Faculty of Letters, Karadeniz Technical University, Trabzon, Turkey
ozbay@ktu.edu.tr

Tuncer AYDEMİR
Faculty of Letters, Karadeniz Technical University, Trabzon, Turkey
tunceraydemir28@gmail.com

ABSTRACT
It is an old consensus by now that a computerized corpus-based analysis of semantic prosodic profiles in English language provides various explanations for how EFL learners exploit the target language during their learning process in terms of lexical selection. The prosodic analysis of lexical selections of EFL learners is likely to give us much better means of understanding the acquisition and learning processes of EFL learners as well as the factors influencing these processes (Granger, 2008). This study investigated the semantic prosodic awareness of Turkish EFL learners in terms of using pragmatic markers such as "I mean, I think, so, well, you know, so on, or so". The analysis was done by comparing two argumentative academic corpora, KTUCALE (Karadeniz Technical University Corpus of Academic Learner English) and BAWE (British Academic Written English). In order to measure possible overuses and underuses as well as statistical significance, frequency counts and log-likelihood ratios were used. Corpus linguistics software AntConc 3.2.2 and SketchEngine online corpus software interface were used in order to collect data. The data was analysed quantitatively by comparing frequencies and calculating log-likelihood ratios. The results of this study validated the commonly held view that a limited number of pragmatic features are overly used in the academic argumentative papers of tertiary level Turkish EFL students. Other findings were that BASE, spoken native corpus, corpus included the highest number of pragmatic markers of all the three corpora and BAWE, written native corpus, contained restricted number of markers such as "so" with ample use.

Keywords: pragmatic markers, corpus, frequency, log-likelihood, discourse markers

INTRODUCTION
Corpus linguistics, a computer based study of language, is commonly defined as a field which enables the compilation of language samples that provides authentic data in written and spoken forms for the linguists who seek to analyse and study the language within itself and with the corpora of other languages. It also gives statistical data such as word lists, frequencies, collocates, concordance and key words. Both native corpora used in the study are compared with a non-native EFL learner academic corpus with an aim to compare the pragmatic marker use. Native corpus (reference corpus) acts as a source and provides reference materials and authentic data, while non-native corpus (KTUCALE) is analysed for such various purposes as lexical content, frequency count and key words. In this study, semantic prosodic awareness of tertiary level Turkish EFL learners in terms of using pragmatic markers was, also, investigated through KTUCALE, BAWE and BASE corpora. When the usage of a word gives "an impression of an attitudinal or pragmatic meaning", this is called a semantic prosody (Sinclair J. , 2000, p. 86) and it is "inextricably linked with corpus data as corpus data enlightens the less transparent meanings of the words and phrases" (Stewart, 2010, p. 79). In this respect, semantic prosodic analysis of pragmatic markers was done with the help of corpus data in order to find out their profiles in the case of spoken and written language.

"I mean, I think, So, Well, You know, You see, So on and Or so" are the target pragmatic markers which were analysed within the scope of the study. The reason for the use of this restricted number of markers in the study is that these pragmatic markers are the most frequent ones which were also used by other linguists in their studies. Part of the reason behind this study is to find out whether EFL learners overuse or underuse certain pragmatic markers when compared to the native speaker use, which deserves immediate focus and research. With the assistance of corpus based research, the probable variations in the use of pragmatic markers can be investigated and analysed in an attempt to raise the awareness of EFL learners towards the use of these items in terms of their syntactic and semantic and pragmatic functions.

Copyright © The Turkish Online Journal of Educational Technology

Semantic Prosody

Semantic prosody has become one of the important notions in Corpus linguistics in recent years. It is also called as semantic harmony, discourse prosody, pragmatic prosody or semantic association. Sinclair (1991) and Louw (1993, p. 157) described semantic prosody as a "consistent aura of meaning with which a form is imbued by its collocates". In other words, according to them, each word may have a special function in its context, and even some near-synonym words may not be used interchangeably. It is also described as "the spreading of connotational colouring beyond single word boundaries" (Partington, 1998, p. 68). Without semantic prosody, words are just considered as single meanings which may not be suitable for writing or speaking; therefore, the discourse may sound non-native or problematic. In Table 1 below, the contextual usages of the pragmatic markers were shown with the examples randomly chosen from BASE, BAWE and KTUCALE corpora.

MARKERS	CORPORA	SAMPLE SENTENCES
I mean	BASE	But **I mean** there it's really striking similarities…
	BAWE	By autocratic, **I mean** that I am solely responsible for the decision…
	KTUCALE	**I mean** that students should do some activities or…
I think	BASE	**I think** that's what he's dealing with there that…
	BAWE	**I think** the pace of the film is interesting.
	KTUCALE	**I think**, a computer cannot take place of teacher.
So	BASE	**So**, it's not an absurd suggestion at all.
	BAWE	**So**, why has this occurred?
	KTUCALE	**So**, I believe that listening helps and facilitates…
Well	BASE	**Well**, he allowed like the Spartans to get help.
	BAWE	**Well**, dear Amanda, thou art the most constant wife…
	KTUCALE	**Well**, in this essay, I will reply how to teach…
You know	BASE	**You know**, Hogarth's last print is called Bathos.
	BAWE	As **you know**, you have been experiencing some chest pain.
	KTUCALE	**You know**, I also use computer program before we…
You see	BASE	**You see**, how this fits here do you see how Spenser…
	BAWE	**You see** he knows me!
	KTUCALE	As **you see**, the world's biggest authorities highlight the…
So on	BASE	We can talk about the viability of it and **so on** but I'm always…
	BAWE	It is quite similar to Burke's; the sea, the storm and **so on**.
	KTUCALE	What their needs and objectives are, and **so on**.
Or so	BASE	Someone asks me to give a talk in a year **or so**…
	BAWE	Over the last half-century **or so**, there has been dramatic decline…
	KTUCALE	Increasingly, in one of the 800 **or so**…

Table 1*:* Contextual usage of the pragmatic markers in BASE, BAWE and KTUCALE corpora

As seen in Table 1, several examples randomly chosen are given in order to exemplify the use of the pragmatic markers across the three corpora. The syntactic positions within the sentence, intended meanings and the semantic prosodic features of the sample sentences above look very similar to each other. Despite the fact that restricted contextual information is provided for each sentence above, it is still possible to speculate that they display similar properties in terms of form and function. Especially, in the example of "I think", it is clearly seen that the marker was used in sentence-initial position in all the examples with the same grammatical function of a noun clause sentence. This and other similarities and possible semantic variations in the use of pragmatic markers across the three corpora were investigated more in detail in results and discussion sections.

Pragmatic Markers

In very general terms, pragmatic markers "presuppose one speaker and at least one addressee taking part in a speech situation, which they at the same time create and monitor via discourse" (Erman, 2001, p. 1337). The various features of pragmatic markers in syntax can be listed as "marking various kinds of boundaries, assisting in turn-taking in oral discourse, expressing speaker attitude and achieving intimacy between speaker and

addressee" (Brinton, 1996, p. 35). Pragmatic markers, such as *you know, I mean, so, well, like* and *now*, have been studied largely from 1980 onwards after corpus based methodologies became common and useful for this kind of studies with the use of computers with the technological developments. They can be defined as linguistic elements "which do not contribute to the propositional content of the utterance which they modify [and that] are frequent in conversation, where they express the speaker's attitudes to the addressee, negotiate background assumptions, express emotions and contribute to coherence" (Aijmer, 2003, p. 1123). According to Jucker (1998, p. 2) who listed the characteristics of pragmatic markers in their study, they are "short, phonologically reduced, optional, multifunctional, and a feature oral rather than written discourse". Moreover, according to Fraser (1996, p. 167), "any signal that has an effect at the communicative as opposed to the strictly propositional level can be considered a pragmatic marker". When pragmatic markers are omitted in a sentence, one result would be that the discourse created in the text becomes unacceptable in grammatical terms and would be seen as "unnatural", "awkward", "disjointed", "impolite", "unfriendly", or "dogmatic" in a communicative context (Brinton, 1996, p. 29). Creating such a discourse could be "incomprehensible for the listener and mission impossible for the speaker" (Svartvik, 1985, p. 352). He also continues that as a consequence of their semantic shallowness, they are difficult to translate into other languages. They can be considered as cultural expressions in oral discourse. As a result, it can be inferred from the previous studies that pragmatic markers play a fundamental role in spoken language.

Research Questions

1. Are the pragmatic markers used mostly in spoken language or written language based on the corpora under examination?
2. To what extent do the tertiary level EFL learners use pragmatic markers while writing?
3. Do tertiary level EFL learners *overuse* pragmatic markers compared to their native counterparts in written and spoken language?
4. Do tertiary level EFL learners *underuse* pragmatic markers compared to their native counterparts in written and spoken language?

METHODOLOGY

The instruments used in the study include two different academic written and one spoken corpora; KTUCALE (Karadeniz Technical University Corpus of Academic Learner English), BAWE (British Academic Written English), BASE (British Academic Spoken English). Moreover, two concordance tools were used in order to obtain and analyse the data, these being Sketch Engine online corpus interface, and AntConc 3.2.2 offline corpus software.

Representation	Corpus	Number of Texts	Average Length of Texts	Total Number of Words
Learner Writing	KTUCALE	196 texts	2,272	509,464 words
Native Expert Writing	BAWE	2897 texts	2,554	6,506,995 words
Native Expert Speaking	BASE	160 texts	7,826	1,252,256 words

Table 2: Corpora Contents of KTUCALE, BAWE and BASE

The non-native learner corpus comes from Karadeniz Technical University Corpus of Academic Learner English (KTUCALE). KTUCALE corpus contains argumentative academic essays which were produced by tertiary level EFL students in Turkey. All the essays are academic in character and the selected sample for the present comparative study is a total of 509.464 words. The native and reference corpus of similar academic content was taken from the British Academic Written English (BAWE) database. This native speaker corpus consists of academic essays written by British university students and contains 6.506.995 words with the contents ranging from Arts and Humanities, Social Sciences, Life Sciences to Physical Sciences in three levels of study: undergraduate, graduate and master levels. KTUCALE and BAWE corpora were analysed via AntConc 3.2.2 offline corpus software which is available free online, and it makes possible to gather data such as wordlist, concordance, collocates, frequency and word clusters, etc... As the final corpus, British Academic Spoken English (BASE) was used in order to compare and to find out whether they are mostly used in spoken or written language. BASE was reached via SketchEngine online corpus interface though which it is possible to gather data from online corpora in various languages, and it is, also, available to upload the researchers` personal corpora for comparison.

In order to compare the three corpora, normalized (standardized) frequencies were calculated. The reason for the use of normalized frequencies is that the raw frequencies do not give the proportional data while comparing at least two corpora because of the difference between the word counts of the corpora. In order to calculate the

normalized frequencies, the following formula was used; (Standardized Frequency = Raw Frequency x 1.000.000 / Corpora Content). Normalizing the data to one million is optional; furthermore, it can be also normalized to one thousand or one hundred thousand. Moreover, LL (Log Likelihood) scores of each pragmatic marker were calculated in order to find out the difference between the use of pragmatic markers in the native spoken, native written, and non-native written corpora. LL scores were automatically calculated via online interface of Lanchester University database on the following link; *(http://ucrel.lancs.ac.uk/llwizard.html)*. If the LL score is over 3.84, critical value is 95th percentile; if the LL score is over 15.13, critical value is 99th percentile.

RESULTS

The pragmatic markers, and their raw and normalized frequencies of the three corpora, which are used in the scope of the study are listed below in Table 3. Two of these corpora are academic written (BAWE and KTUCALE) and the other (BASE) is academic spoken corpus.

	BASE	BAWE	KTUCALE	BASE	BAWE	KTUCALE
		Raw			Normalized Per Mil	
I mean	1052	27	13	840,08	4,1493	25,5170
I think	1501	318	122	1198,60	48,8704	239,4673
So	12790	10478	901	10213,60	1610,2671	1768,5253
Well	3187	50	10	2545,01	7,6840	19,6284
You know	2133	50	37	1703,30	7,6840	72,6253
You see	381	24	9	304,25	3,6883	17,6656
So on	385	145	44	307,45	22,2837	86,3652
Or so	74	34	3	59,09	5,2251	5,8885
TOTAL				**17161**	**1710**	**2236**

Table 3: Raw and Normalized per mil frequencies of the pragmatic markers in three corpora

According to Table 3, it is seen in the total normalized frequencies of the three corpora that target pragmatic forms are overly-used in spoken corpus. These overused-pragmatic markers seem natural if we consider the nature of spoken language and how much these markers fit into it. In other words, pragmatic markers have speech-like nature and EFL learners are quite expected to use them in spoken discourse. With all these in mind, however, in the other two academic corpora there are also significant differences in terms of the usage of these markers as well. For instance, based on the normalized frequencies above, it seems that the KTUCALE corpus display higher usages of pragmatic markers (2236) when compared to BAWE (1710). This is most visible with the exception of "so" and "or so" whose normalized frequencies are almost the same. The reason for doing so may be given to the fact that Turkish EFL academic writers overused these speech-like items in their academic essays unlike their native speaker counterparts who seemed to have used much fewer of these markers. With all these, however, there is still a need for a statistical measurement in order to demonstrate the difference between the groups. Table 4 and 5 below show the LL measurements for the overall frequencies among the three corpora.

	BASE (normalized)	BAWE (normalized)	LL score
I mean	840	4	1119.24
I think	1199	49	1316.77
So	10214	1610	6981.11
Well	2545	8	3430.99
You know	1703	8	2270.14
You see	304	4	384.28
So on	307	22	294.58
Or so	59	5	53.63

Table 4: Log likelihood Score of BASE and BAWE

The first comparison was made between the two native reference corpora. In Table 4 above, log likelihood scores of BASE and BAWE were calculated in order to test the commonly-held view that pragmatic markers are mostly used in spoken language as compared to written one. As seen in Table 4, all of the LL scores are highly over 15.13 ($p<0.01$) which is the critical difference value of 99.99th percentile level. It is clearly seen that there is a significant difference between BASE and BAWE corpora regarding the pragmatic marker use, which means that the target pragmatic markers can be considered as the speech-like words and phrases with their oral functions.

With this comparison, on the other hand, it is also seen that BAWE corpus include few instances of the pragmatic markers with the exception of "so", which signals the fact that these forms may not be ideal forms to be used in the written academic language by the native writers. This contention can further be increased by the very few uses of pragmatic markers such as "I mean (4)", "well (8)", "you know (8)", "you see (4)", "or so (5)" in BAWE. In BASE corpus, however, it is seen that there are ample amount of pragmatic marker use with "so" the most (10214). This is followed by "well" and "you know" markers which have a normalized value of total 4250 per million. "I mean", "you see", "so on" are the other types of pragmatic markers which received higher values when compared to their native written corpus hits.

	BAWE (normalized)	KTUCALE(normalized)	LL score
I mean	4	26	18.03
I think	49	239	136.54
So	1610	1769	7.48
Well	8	20	5.31
You know	8	73	60.07
You see	4	18	9.64
So on	22	86	40.53
Or so	5	6	0.09

Table 5: Log Likelihood Score of BAWE and KTUCALE

In Table 5, log likelihood scores of BAWE and KTUCALE were calculated in order to compare the pragmatic marker content of the native academic corpus with a learner academic corpus of the tertiary level Turkish EFL learners. According to the Table 5 above, all LL scores except *or so* are over 3.84 ($p<0.05$) which is the critical difference value of 95 %. When compared with the native reference corpus, the most striking finding in Table 4 is that Turkish academic corpus seem to include highly significantly overused markers with the exception of "or so". The highest overuse in KTUCALE corpus is with "I think", "you know" and "so on" markers. One reasonable explanation for the overuse of pragmatic markers in the KTUCALE may be the language transfer from the L1. The learners` native language also contains similar markers and it is possible that the learners may have transferred them directly from L1.

	BASE (normalized)	KTUCALE (normalized)	LL score
I mean	840	26	967.02
I think	1199	239	699.83
So	10214	1769	6580.55
Well	2545	20	3321.84
You know	1703	73	1853.12
You see	304	18	307.58
So on	307	86	131.83
Or so	59	6	50.09

$p < 0.01$ (critical value: 3.84); + indicates overuse in the first corpus relative to the second corpus; -indicates underuse in the first corpus relative to the second corpus

Table 6: Log Likelihood Score of BASE and KTUCALE

In Table 6, log likelihood scores of BASE and KTUCALE were calculated in order to compare the pragmatic marker content of the native academic spoken corpus with KTUCALE. According to the Table 6 above, the log likelihood scores of BASE and KTUCALE were calculated with an aim to compare native spoken corpus with a non-native written corpus in order to observe the probable underuse and overuses in the non-native learner corpus. As seen in Table 6, all of the LL scores are highly over 15.13 ($p<0.01$) which is the critical difference value of 99.99th percentile level. It is clearly seen that there is a huge difference between BASE and KTUCALE corpora regarding the pragmatic marker use. The similar difference was also observed with the BAWE corpus as well. The fact that written academic corpora, regardless of native or non-native, contain much fewer pragmatic markers than the spoken component once more support the long-validated view that pragmatic markers are very closely related to the speech.

In Table 7 below, semantic prosodic profiles of all the pragmatic markers were given across the three corpora. Some interesting conclusions can be drawn from these prosodic profiles. The positive, negative or neutral features of each marker were not indicated in this study. Instead, each pragmatic marker was analysed with only their 1+ right or 1+ left complements. First of all, Table 7 indicates the results of the target marker "I mean". The most frequently used 1R adjective is the subject pronoun “I” (140 times) in BASE and clausal “that” in BAWE (5 times) and KTUCALE (5 times). The second most frequently used collocate is “It” in BASE (105) and this item was also used in the other corpora a total of 3 times. It seems that “I” is a significant collocate of the pragmatic marker “I mean” in BASE but BAWE and KTUCALE corpora do not contain this item. On the other hand, “it” is another significant collocation in BASE (105 times) but the two written academic native and non-native corpora contain very few instances of this item after “I mean”.

The pragmatic marker “I think” displays more or less a similar picture in the three corpora with the difference that absolute frequencies are higher in each. “It”, “That” and “I” complementation were used after “I think” marker in an overwhelming manner in BASE (506 times). However, the same complementation “that” and “it” were used only a total of 108 times in BAWE and 48 times in KTUCALE after “I think” marker.

The pragmatic markers “So”, “Well” and “You know” also display much higher frequencies in BASE but much lower frequencies in BAWE and KTUCALE corpora. Based on Table 7, it seems that “You” complement was used 1777 times with pragmatic markers “So”, “Well” and “You know” in BASE, but only 2 times in the academic written corpora.

Another interesting finding is that “You see” marker received different complements across the three corpora. The variety of complements with this marker may not be easy to explain. One thing is certain, however, that it received very few complements in BAWE and KTUCALE (a total of 17) unlike BASE corpus, which received more than 100 complements.

	BASE (Abs)	**BAWE (Abs)**	**KTUCALE (Abs)**
I mean	I (140) It (105) There (55)	That (5) When (1) It (1)	That (5) My (2) It (2)
I think	It (233) That (163) I (110)	That (67) It (41) This (23)	That (43) This (11) It (5)
So	You (1432) We (1291) That (1259)	It (464) We (223) They (196)	They (67) It (45) Students (17)
Well	I (205) It (160) You (128)	I (5) Let (3) Now (2)	You (2) Then (1) Helena (1)
You know	You (217) It (140) That (131)	I (4) That (3) Your (1)	That (5) I (2) How (2)
You see	That (51) This (29) I (25)	We (2) My (2) It (2)	Words (4) Result (1) It (1)
So on (**1L**)	And (359) It (7) This (4)	And (128) Skills (3) Time (2)	And (42) Books (4) Languages (2)
Or so (**1L**)	Years (19) Minutes (6) Fifty (6)	Last (5) Decade (5) Years (3)	Years (1) Past (1) -

Table 7: Semantic profiles of all the pragmatic markers across the three corpora

Finally, since “So on” and “Or so” pragmatic markers were in sentence-ending positions their 1Left collocates were listed. It is interesting to note that “and” complement was used more than 350 times in BASE corpus and very few instances of “it” and “this” complements followed (11 times) this. The number of “and” complement in BAWE corpus also deserves further attention (128 times). However, as a non-native corpus, KTUCALE corpus includes “and” complement only 42 times, much fewer than that of BAWE. Nevertheless, KTUCALE

corpus "So on" marker is also followed by very few nouns such as *books* and *languages*. As a result, considering the contents of all three corpora, it is possible to say that the distribution of the profiles have no similar proportions to each other when these pragmatic markers are in the sentence ending positions.

Discussion and Conclusion

This study focused on the contrasts between native and non-native performance in academic writing in terms of pragmatic markers and is likely to give us a more precise understanding of the challenges that tertiary level EFL writers encounter while writing academic essays and offers a potential basis for pedagogically useful conclusions.

Considering the normalized frequencies and log-likelihood results of all the three corpora used in the study, it is possible, first of all, to conclude that BASE corpus included the highest number of pragmatic markers. This is, of course, due to the fact that pragmatic markers are generally regarded as speech like items and the BASE corpus was a spoken academic corpora. Especially, such markers as "So" (10213,60 occurrences per million words), "Well" (2545 occurrences per million words) and "You know" (1703,30 occurrences per million words) became the most widely used ones through their pragmatic meaning and function in communication.

However, the pragmatic marker use in the two written academic corpora deserves further pedagogic analysis and interpretation. Based on Table 3 above, it is interesting to note that Turkish EFL learners when compared to native English writers (BAWE) overuse almost all the pragmatic markers. The reason for these overly used pragmatic markers may be given to the several related factors. In the literature, similar results have also been obtained in various studies such as Gilquin & Paquot (2007); Simcikaite, (2012); Babanoğlu (2014). As stated by Gilquin & Paquot (2007) and by Babanoğlu (2014), majority of the problems in the use of pragmatic markers by Turkish EFL learners may be given to the various factors such as L1 transfer, L2 instruction, register confusion and developmental factors.

From a language learning and teaching perspective, first of all, we believe that there is certainly a need to raise the tertiary level EFL students` awareness towards the use of pragmatic markers since they overuse some markers and underuse others. This needs to be balanced and, therefore, awareness raising activities are needed in the language classrooms. These awareness raising activities need to be targeted toward the use of pragmatic markers and tertiary level EFL teachers need to draw their attentions of their students towards the correct use of these markers in writing. In other words, considering the fact that EFL learners lack competence in terms of using pragmatic markers in their academic essays, there is a need for integrating the use of pragmatic markers into the syllabus and they should be able to know the differences between speech like items and academic English.

It is also the case that L1 transfer may have played a role for the overly- used pragmatic markers in the essays of Turkish EFL learners. In fact L1 negative transfer is a common phenomenon for many EFL learners, which may also account for the Turkish learners' overuse of oral features in their writing.

Finally, considering the commonly held view that Turkish EFL learners experience difficulties learning the English language despite devoting a notoriously extended period of school time learning it, the over use of pragmatic markers in the academic essays of Turkish EFL learners may be attributed to the interlanguage developmental levels which are formed by the learners of a foreign language (L2) who have no native-like proficiency but approximate the target language and preserve several elements of their L1 while writing in L2.

Last but not least, the understanding of the developmental stages of the EFL learners should be accepted as a potential advantage of the Corpus based Interlanguage analysis process (Meunier, 2002). It is through these kinds of contrastive interlanguage analysis procedures that the problems of EFL learners are revealed.

REFERENCES

Aijmer, K. &.-M.-V. (2003). The discourse particle well and its equivalents in Swedish and Dutch. *Linguistics 41 (6)*, 1123-1161.

Babanoğlu, M. P. (2014). A corpus-based study on the use of pragmatic markers as speech-like features in Turkish EFL learners' argumentative essays. *Procedia - Social and Behavioral Sciences 136*, 186 – 193.

Brinton, L. J. (1996). *Pragmatic Markers in English.* NewYork: Mouton de Gruyter.

Erman, B. (2001). Pragmatic markers revisited with a focus on you know in adult and adolescent talk. *Journal of Pragmatics 33*, 1337-1359.

Fraser, B. (1996). Pragmatic Markers. *Pragmatics 6*, 167-190.

Gilquin, G. &. (2007). Spoken Features in Learner Academic Writing: Identification, Explanation and Solution. *Proceedings of the 4th Corpus Linguistics Conference.* Birmingham: University of Birmingham.

Granger, S. (2008). Learner corpora in foreign language education. In N. V.-S. Hornberger, *Encyclopedia of learner and education. Second and foreign language education, 4* (pp. 337-51).

Jucker, A. H. (1998). Discourse markers: an introduction. In A. H. Jucker, *Discourse markers. Descriptions and theory* (pp. 1-12). Amsterdam: Benjamins.

Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. &. Baker, *Text and Technology.* Amsterdam: John Benjamin.

Meunier, F. (2002). The Pedagogical Value of Native and Learner Corpora in EFL Grammar Teaching. In S. Granger, & J. &.-T. Hung, *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 119-142). Amsterdam: John Benjamins.

Partington, A. (1998). *Patterns and Meanings: Usig Corpora for English Language Research and Teaching.* Amsterdam: John Benjamins.

Simcikaite, A. (2012). Spoken Discourse Markers in Learner Academic Writing. *Studies about Languages, 20*, 27-34.

Sinclair, J. (1991). *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.

Sinclair, J. (2000). Lexical Grammar. In *Naujoji Metodologija* (p. 86). Darbai ir Dienos.

Stewart, D. (2010). *Semantic Prosody: A Critical Evaluation.* London: Routledge.

Svartvik, J. &.-B. (1985). Words, words, words: The rest is silence. In S. &. Bäckman.

View publication stats