

Juliana ÇYFEKU

University of Korça, Albania

juliana.cyfeku@gmail.com

Lindita KAÇANI

University of Korça, Albania

lkacani@yahoo.com

A VIEW OF CROSS-LINGUISTIC DIFFERENCES IN ENGLISH LANGUAGE CORPORA; AN INTER-LANGUAGE ANALYSIS TO THE TARGET LEARNER

Pursuant to the fact that native English speakers differ widely from foreign English speakers in terms of communicative aspects, this study provides essential observational remarks elaborated from the researcher viewing English language acquisition from a communicative perspective. Methodologically, it attempts to relate the communicative challenges hereby referring to the ‘miscommunication occurrences’ that university students encounter in the English language learning process. The study develops with an emphasis on elaborating communication in English vocational language learning in reference corpora use.

Language corpora display and furnish university students with different communicative perspectives and guide them to a variety of data-driven inquiry for certain linguistic purposes. Indeed, the flow of English language authentication elicited from corpora exploitation highlights the cross-linguistic differences that target students find difficult to comprehend in the English language learning process. English language corpora offer a range of methodological tools that enable learners to conduct any inter-language analysis to gain communicative insights in real life contexts. Throughout, the implementation of corpus-based analysis students can autonomously investigate and draw comparisons upon almost any language patterns in terms of lexical, structural or lexicon-grammatical occurrences. This permits them to contrast those cross-linguistic native-like language differences they notice in the plethora of corpora language. The final outcomes would no doubt yield in crafting the target learners’ communicative competence providing ‘in between the corpora lines’, clues for self-correction of any possible individual creation irregularities.

Key words: communication, occurrence, cross-linguistic difference, corpora, inter-language analysis

1. Introduction

The improved accessibility of computers has revolutionized the study of language especially in second language acquisition over the last few decades. The main purpose of conducting this study is to shed light on the authentic English language occurrences that corpora offer. It also points out the fact that the study of

corpora has changed from a subject for specialists to a field of inquiry resulting to be open to all kinds of learners, teachers and researchers.

Applying corpus language for linguistic issues in foreign/second acquisition enhances the teaching and learning process. This study gives prominence on the effect of corpus studies upon theories of language and how languages should be described. Corpora allow all kinds of ‘researchers’ to count categories in traditional approaches to language, draw comparisons among authentic usages of L2 as displayed in them and foster them to observe and investigate patterns of L2 that were not noticed before in speech. Moreover, this study intends to raise the L2 learners’ interests in how language, more specifically English, works, and how knowledge about it can be applied in certain real-life contexts. The reader will be informed on the methodological tools that can be easily found online for specific classroom use. Evidently, corpora application allows language learners to carry out corpus investigations for linguistic purposes in SLA.

Provided the large number of examples of a word or phrase which is presented in an electronic form, corpora offer obvious didactic advantages for data-driven analysis. The corpus is usually subjected to a concordance analysis by using analytical software (a concordancer) to find the occurrences of a particular word or phrase in a text or in a compilation of texts. It presents to teachers the possibilities in designing English language learning tasks. This serves the need to achieve an adequate proficiency in English language provided that the latter is currently positioned as the primary language of international communication in almost every field of professional activity. Specifically we embrace what Widdowson’s (2000) statement [...that correct English language use in oral and written discourse requires the possession of a solid linguistic competence which includes the knowledge of grammar rules and vocabulary items, as well as the ability to use them appropriately in real-life contexts].

The importance of the development of linguistic competence, being a component of communicative competence, has been the main aspect of inquiry considered as the pitfall of the “communicative approach” to foreign language teaching (Widdowson, 1998, 1990; Sheils, 1993). In addition other researchers in this domain of inquiry, such as Long & Robinson (1998), and Pearson (1996); have highlighted the important role of grammatical correctness in foreign language communication. Communication in EFL seems to be the gap which learners find hard to overcome as not feeling furnished with the proper communicative competence. Purposefully, what these researchers argued was that learners could achieve communicative competence in English by conducting linguistic analysis in terms of form and functions as it authentically used in real life. On such a basis foreign English learners can make comparisons in terms of linguistic irregularities of lexical use from their part, thus allowing room for self-awareness and self-correction. Indeed, this has been one of the arguments in favour of the use of a linguistic corpus as an invaluable source of natural examples of language use.

Specifically, a corpus is a collection of authentic texts (written or spoken transcripts) that are stored in an electronic form (Hunston, 2002: 2). It is defined in terms of both its form and its purpose. Linguists have used the word *corpus* to describe every occurring pattern, phrase, single sentence to a set of written texts or tape recordings, which have been collected for linguistic study. Because computers can hold and process large amounts of information, electronic corpora are easily found and used to study aspects of language. Since these are usually larger than the small, paper-based collections that we find in any EFL course book or other similar sources, the storage of the texts in such a form allows their compilation in large quantities, amounting to millions of words, which, as Biber, Conrad & Reppen (1998) pointed out, make these texts available for both qualitative and quantitative analysis either from the researchers or the learners themselves. A corpus may be composed of texts from various sources (e.g., a reference corpus such as the British National Corpus or the Bank of English), or from a particular genre (e.g., a specialist corpus). The corpus is designed and stored in such a way that it can be studied non-linearly, and both quantitative and qualitative linguistic analysis can be obtained. This differentiates a corpus from a library or an electronic archive. The purpose is not simply to access texts in order to read them for linguistic information but to investigate beyond the lines to figure out the irregularities of language as used from the native speakers and/or writers in authentic contexts. Since they refer to genuine discourse, there are no strict rules or specific ways about how large a corpus needs to be, but it is generally assumed that it should provide enough examples of language to ensure relevant generalizations on language form and functions (Baker, 2006).

Due to the wide range of corpora and based on the specific and linguistic demands prevailing the language needs analysis, corpora can be analyzed traditionally or in a modernized way through computerized tools. Depending on what we need to analyze, we can observe and calculate the occurrences of linguistic terms in text samples, or we can analyze them in a more rapid way with the aid of corpus analysis tools, such as word lists and concordancing software. Specific information on the ways of operating and utilizing corpora data provided with practical examples in SLA can be found in the corpora studies of Bowker & Pearson (2002), Biber, Conrad & Reppen (1998); and Tribble (2002). The easiest method to explore corpora is by means of concordance. A concordance (as can be seen from Fig. 1), is an electronically list of all the occurrences of a particular word or phrase in a corpus. It is non-ordinary presented within the context they occur in (usually a few words to the left and right of this word) which at first glance seem ambiguous and irrelevant to the sequenced phrases as displayed in the lines. A concordance is sometimes called a “key word in context” (KWIC), which means a word or phrase that is of interest, for a researcher or a teacher or both (Baker, 2006: 71).

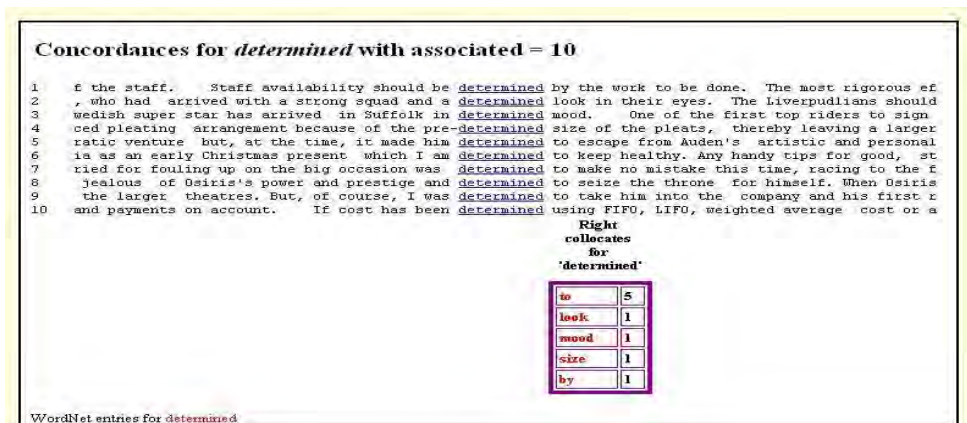


Figure 1. Online Concordancer output

Amongst the pioneer teachers of the concordance use, we mention Tim Johns, who was the author of the concept of Data Driven Learning (DDL, 1991). Various corpora studies shed light on the use of concordances software and their implementation in real classroom contexts. Relying on such novelty use from the native speakers (English as L1) as corpora furnish with a plethora of genuine use the student learners are simultaneously acquainted with corpora and its linguistic significances in L2 (English viewed as L2). It promotes DDL as an approach to language learning based on the assumption that the use of authentic language together with a concordancer will enable the learners to gain insights into the language as it is used in real-life situations. Indeed DDL refers to the learning process which is based on the learner’s discovery of rules and patterns of language use, and not on the teacher’s explanation and provision of examples of the language invented for pedagogical purposes. The teacher presents oneself in the role of an instructor in case of requirement from the autonomous learners.

Moreover, the role of corpora in language pedagogy has received considerable attention in the literature concerned with the teaching of languages. A corpus does not contain new information about language, but the software used for linguistic investigation offers a new perspective on the familiar. Specifically, software packages by means of various tools process data from a corpus in three ways: showing frequency, phraseology, and collocation of L2 occurrences. For instance the interpretation of concordance lines can facilitate developing linguistic competence. Issues of these kinds provide genuine occurrences of language and evidently have been dealt with practical evidence for solutions by a number of researchers such as: Baker, 2006; O’Keeffe, McCarthy & Carter, 2007. Relevantly, language corpora have been viewed as a facilitating and contemporary aid in the understanding of particular text types, genres and styles in addition to the designing of syllabi and materials for foreign language courses in reference corpora collection of native linguistic occurrences (Biber, Conrad, & Reppen, 1998; Tribble & Jones, 1997).

2. A VIEW OF CROSS-LINGUISTIC DIFFERENCES THROUGH LANGUAGE CORPORA

Native English speakers differ widely from target English speakers in terms of communicative aspects either in written and/or oral discourse. This study provides a personal view of English language acquisition from a communicative perspective due to the cross-linguistic differences the student learners encounter in English language learning. It attempts to relate the communicative handicaps that mostly ESP students encounter in their vocational learning process at university level. Pursuant to the fact that English results as a compulsory academic subject the study develops to contemplate the academic needs to reach to perfection on the basis of vocational language learning. On purpose it observes and examines the degree to which the frequency and co-occurrence of corpora vocabulary use leads the students to DDL and contributes to better language performance. The latter, resulting so with an elaborated communicative competence of overall English students targeted to ESP courses.

Basically, relying upon thorough language course observations the results obtained may report such feasible affective dependencies among learners and language course needs as determined in our syllabi. Due to the different communicative perspectives and the variety of data-driven inquiry that language corpora display, English language instructors should adopt various communicative contexts in EFL teaching. That is indeed the case to measure the affective *variables* in English language acquisition and learning and to put clear cut ends so as to prevent any possible linguistic miscommunication from the students in authentic situations. Provided the reliable linguistic differences the students notice when exploiting corpora such linguistic ambiguities are relevantly clarified to any language inconvenience. In addition, they are motivated and corpus-guided to gain control and exert their communicative skills over language.

In dependence to the language proficiency that the learners possess, a language corpus may be greatly helpful for further scientific learner linguistic analysis. The most frequently used method to analyze corpora is the contrastive inter-language analysis that was established to serve the purpose of corpus research by Gragner (1996). Its implementation in the research provides clues to demanding learners to draw comparisons; speculate on the words' behavior through the lines; and formulate rules of grammar among the concordance lines of the key words in the context that constitute authentic occurrences in reference situational happenings expressed by natives in the English language. Thus they may serve as challenging out of traditional linguistic samples for authentic classroom language implications. Irregularities of language lexicon make the learners think critically, logically and that assists in reinforcing best what is uncommon in language. Basically this approach enables learners to view the language differences that exist between native speaker data and foreign language data (L2 vs. L1) or even among the learners themselves (L2 vs. L2). It displays corporal methodological features of qualitative and quantitative differences in the field of EFL/ESL and mainly ESP vocabulary acquisition. This type of analysis is greatly facilitated by computational programs

that are huge in quantity but pointing out a contemporary one such as *Compleat Lextutor Text Tools* (Cobb, 2000). As displayed in the Figure. 2, it processes words of a certain language in any required use or function in discourse. It gives researchers, language practitioners and learners immediate access to the words or phrases that are significantly un/frequently used, overused or misused by learners. It actually finds out and makes evident all the possible occurrences of the language items in question aimed for linguistic examinations. Applying its computational methodological tools makes relevant the simultaneous procession of a range of texts in written or transcript form. Small and bigger differences in frequency of vocabulary use can be a matter of linguistic variance among learners and therefore they need to be interpreted prudently.

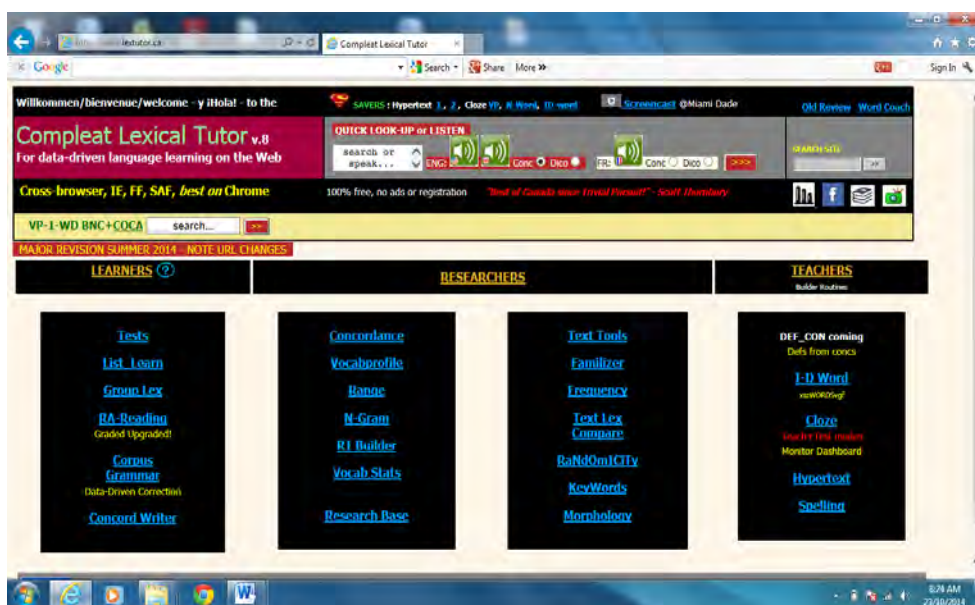


Figure 2. The Compleat Lexical Tutor, site entry web page

A corpus-based analysis can investigate almost any language patterns in terms of lexical, structural, lexico-grammatical occurrences or individual creation irregularities as discovered from and intended to distinguish native *versus* target usage. Choosing the proper analytical tools, as overstated for instance Cobb's Lextutor (2000), we can discover not only the patterns of the learners' language use, but even the extent to which they are used in relevance to the contextual factors that influence variability in and between them. Indeed the frequency word lists allocate the various English situational vocabularies each group of inquiry (hereby pertaining to various social groups) had employed in the discourse. The overview as presented from the various corpus researches show considerable variation in the numbers of subjects (large-scale or small-scale), in the backgrounds they had (evolving socio-

cultural factors), and in the empirical data gathered for elucidating outcomes, which come from corpus collections such as tape-recorded samples of speech, from student writings, from various types of tests and corpus-based tasks (Remarks withdrawn from Çyfeke (2014); a corpus-based study conducted with EFL students). Text analyzers fulfill the analytical function of the students' tasks to obtain elaborate evaluation at the end of their L2 learning and overall teaching process. Thus for instance, once students' texts or written tasks have been generated in a word processing file, text analysis applications can research and examine documents for various features, count them, and give student writers or teachers quantitative information about the distribution of textual features. Most word processing packages include spelling, grammar, and style checkers. They also include counters, which can tabulate the number of words in a text, the average word or sentence length, punctuation marks and so on. Beyond the commercially available features of word processors, more specialized programs can tag and count various parts of speech and lexical features (Biber, Conrad & Reppen, 1998). Some studies have suggested that particular types and combinations of syntactic and lexical features in student texts may be positively correlated with text quality or L2 proficiency. Procedurally, the implementation of learner corpora assists any kind of researcher to notice and analyze: the learners' language use (formal, informal or a highly specific targeted register), the different meaning and the potential of the words as displayed in the corpora, the use of the typical specialized terms targeted for comprehension. Certain language features are more typical for different types of learners such as: less or more proficient; motivating or demotivating due to the lack of the cross-linguistic or socio-cultural differences; or the ones who are curious to put words together for a better linguistic comprehension. Thus under the effect of corpora involvement they immensely try to get the utmost from the language usage and challenge the native's capacity to hinder fast and indirect communication. And to support our point, a careful planning in our syllabi evolving a deliberate combination of methods may serve the purposes of a foreign language classroom needs analysis research. Indeed, this is precisely what language teachers and students are encouraged to do so as to yield to practicality and validity of their findings and demonstrate their methodological contribution to the field of SLA.

3. RESULTS AND CONCLUSIONS

The examining corpora effect and outcomes were firstly related in English for Specific Purposes language learning. This new trend in SLA gave rise to Corpus linguistics methodology with the advent of powerful personal computers at the end of the 1980s. It has assisted foreign language research to resolve specific linguistic issues. It fosters the intent to the demanding learners to conduct various inter-language analyses on linguistic issues that seem ambiguous or cause irregularities in the perception of the cross-linguistic occurrences in English. Corpus linguistics (defined in Reppen, 2010) is ideally used to investigate whether a particular word or phrase is used more frequently in either formal or informal style, in specific genres by English speakers. Applying it for linguistic issues in foreign/second

language acquisition enhances the teaching and learning process. It offers a range of methodological tools that facilitate the comprehension especially on genuine vocabulary presentation and provides clues for data-driven learning. The sample website entry in Figure. 3 reveals specific data such as the frequency of certain words which is unusually high in comparison with some norms of the words' performance in corpus contexts. Given its potential as a powerful approach in SLA the researchers as well as learners can make claims on certain issues which other approaches cannot offer even in the case of possible solutions for the same concerning matters.

Rank	Freq	Keyness	Keyword
1	154	558.668	company
2	71	496.393	Ness
3	59	412.496	Google
4	44	307.624	Newmont
5	103	267.768	says
6	49	239.526	executive
7	34	237.709	Bechtel
8	30	209.744	CEO
9	29	202.752	Buyat
10	28	195.761	PR
11	27	188.769	paroxetine
12	26	181.778	Ketchum
13	24	167.795	http
14	24	167.795	Mart
15	24	167.795	www
16	24	159.459	Wal
17	31	157.437	corporate

Figure 3. A keyword list of a business text corpus data elicited from Scott, 2004

Specifically, corpora have been used in combination with speaker-annotated data to investigate language usage. Various researchers have contributed in the field of SLA for instance by examining corpora to be representative of a variety of language. Precisely, Rayson, Leech & Hodges (1997) and Schmid & Fauth (2003) cited in Baker, (2006:75) examined the spoken section (about 10 million words) of the British National Corpus (BNC). In their study they identified their speakers and respondents by the factors of gender, age and social group. They have undertaken selective qualitative analyses to highlight the very specific differences in vocabulary use from males and females as displayed for instance in the British National Corpus. They contrasted frequency lists derived from different sectors of the corpus. They found that female speakers tended to dominate talking in comparison to male speakers. In their research, Rayson and colleagues confirmed the ‘difference’ hypothesis that male speech was more ‘factual’ and concerned with reporting information, meanwhile female speech was more interactive and concerned with establishing and maintaining relationships (the male ‘rapport’ vs. female ‘rapport’ distinction). They assert that common outcomes were previously obtained in other corpus-based studies of spoken and written English. Altogether they concluded

that ‘prepositions’ and the definite article ‘the’ are highly associated with the ‘informative’ and nominal tendency of written language; whereas first – and second-person pronouns are associated with the ‘interactive’ tendency of spoken language. Indeed this fact runs counter in previous studies and is fully evidenced from EFL/ESL students when contrasting corpora language and their topical related written language. Analytically students strive to elicit the differences in discourse from the concordance lines of a specific word they are examining. The keyword selected that constitutes a small corpora in many cases when compared to the precise use from their own in discourse results to contain a much larger representation as more ‘interactive’ than in their oral or written discourse. These remarks are assembled in terms of communicative differentiation between the two cross-linguistic English eloquitors. If their written discourses in reference specific words are statistically processed, the data analysis will for certain reflect a real difference of native *versus* target linguistic behavior in the academic written language. This for instance can apply to compare texts or writings of the same topic by utilizing the “type-token ratio”, a valid measure which indicates the number of different words used across texts referring to types divided by the number of words on which they are based that are tokens (Baker & Ellece, 2011). The following Figure 4. addresses the evidenced homepage for eliciting frequency features to raise students’ awareness on the variety of words used in the corpora and at the mean time assists them increase the variety of words used in their writings.

Home > Frequency

Frequency Home Page

Extract basic frequency lists from user texts or corpora; download influential frequency lists including AWL and Français Fondamental

For related analyses check [Keywords](#) (text vs. general frequency) and [Range](#) (frequency across texts).

Menu

- English
- French
- The Compleat Lister
V2.3: Integration of Freq and N-Gram, Jun 2010
- Frequency Trainer
Quiz Builder (v.2 August 2013: INC+OCCA)
Sort headwords into 1-2k, 3-5k, 6-10k, 11-15k categories
- Famous frequency lists for download

Sample output for 1 + 2

RANK	FREQ	COVERAGE		WORD
		Individual	Cumulative	
1.	11	5.50%	5.50%	THE
2.	9	4.50%	10.00%	BE
3.	6	3.00%	13.00%	FOR
4.	5	2.50%	15.50%	OF
5.	5	2.50%	18.00%	TO
6.	4	2.00%	20.00%	DAY
7.	4	2.00%	22.00%	REMOVED
8.	4	2.00%	24.00%	HE
9.	4	2.00%	26.00%	IN
10.	4	2.00%	28.00%	NP

Eng and Fr Freq programs by Cathy Bailey, Georgetown Linguistics, adapted T Cobb, Compleat Lister and Freq Train by T. Cobb, latter based on Vp-INC lists by P. Nation, VUW-Wellington NZ

Figure 4. Homepage for Frequency features

Additionally, Catalan & Ojeda Alba (2008) in their quantitative study also examine the English vocabulary differences of males and females. Using a small corpus of learner English, they found that on average females produced longer compositions as well as overall gender differences in the type of the vocabulary

they used. Harrington (2008) also emphasizes the use of a careful qualitative analysis to investigate upon students' corpora as the incorporation of corpus linguistics techniques can provide evidence for gender differences individually and no speculations should be made for over generalizations. The analysis of gendered speech styles helps *target* learners to understand certain ways in which native male and female students characteristically tend to use language, and that awareness of such phenomena can be applied in various educational academic contexts such as L2 language settings. Since variations persist even in the natural occurrences of the natives as inferred from the corpus researches, variances in apprehending the communicative meaning from target learners will also be encountered in EFL. Not even two identical twins equally perceive the same phenomena; the stimulus that is the authentic language occurrence may be the same but the logical and individual explanation may vary slightly or may be totally diverse.

Evidently, the present study attempts to give adequate answers and viable solutions to such variance in communication throughout the students' awareness of the English language use of corpora to reinforce EFL/ESL or ESP vocabulary. It presents a range of methodological tools that facilitate the vocational comprehension especially on genuine vocabulary presentation and provides clues for data-driven learning. Throughout, the implementation of corpus-based analysis students can investigate autonomously and draw comparisons upon almost any language patterns in terms of lexical, structural or lexicon-grammatical occurrences. Methodologically, corpora use in the English classroom results as a suitable way of understanding English in authentic discourse and for the most it assists learners crafting their communicative competence. It is the communicative competence the linguistic gap that most learners find frustrating and as a result feel demotivated to delve autonomously especially in the oral discourse. Noticeably, their direct involvement in the plethora of corpora language fosters the desire to get involved in the corpus analysis, to get the utmost from the English native's communicative language. Once dealing with corpora data they individually pursue certain language investigations with a targeted linguistic matter to deal for the sake of differentiating communicative and inter-language usages of English. This furthermore enables them to gain linguistic competence for self-correction of their individual creation irregularities when contrasted to the cross-linguistic differences they notice in reference corpora language.

To sum up, the success in implementing corpora in all field-specific English language auditoriums relies on the fact that the communication competence in English has a strong dependency in the understanding of language in all the possible forms that natives shape it contextually. The understanding and as a result the creation of discourse can be achieved in a satisfactorily manner by the use of any kind of corpus concordancer (hereby referring to a range of online corpus resources) and by the students themselves individually performing various language-learning tasks. Provided the linguistic methodological tools that language software packages offer teachers can facilitate and guide students towards an analysis of the texts included in the corpus. Thus by relating corpora elements in language teaching, all

the components of the target language linguistic research i.e. language teachers and student learners assume to be successful. Considered so as useful and innovative tools (e.g. concordance lines) aiming at the development of students' communicative competence it no doubt results effective for them to formulate a final correct discourse. In this way both sides can have their ends meet in the foreign language learning process and that means overcoming the cross-linguistic differences by mastering contextually the hinted inter-language significance throughout corpus data. Notably once the linguistic occurrences are introduced to the students and respectively guided to autonomous investigation, the outcomes would definitely yield in striving and grasping for a final equalization of written and oral discourse. They will gain competence and feel confident to challenge native communication in English. The whole process resembles with the reverse sides of the shaft with all its infinite numbers (hereby referring to the range of lexicon) that is approximately equalizing competences of target *versus* native in the field of language authenticity and ease of communication in and/or between both kinds of speakers of English.

REFERENCES

- Baker, P. (2006):** *Using corpora in discourse analysis*. London, New York: Continuum.
- Baker, P. & Ellece, S. (2011):** *Key Terms in Discourse Analysis*. New York: Continuum
- Biber, C., Conrad, S., & Reppen, R. (1998):** *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Bowker, L., & Pearson, J. (2002):** *Working with specialized language: A practical guide to using corpora*. London, New York: Routledge.
- Catalan, R M^a J. and Alba, O. J (2008):** *The English Vocabulary of Girls and Boys: Evidence from a Qualitative Study*. In: Harrington, K. Litosseliti, L. Sauntson, H. and Sunderland, J. *Gender and language research methodologies*. UK: Palgrave Macmillan.
- Cobb, T. (2000):** *The Compleat Lexical Tutor* [Website]. Available from: <http://www.lextutor.ca/>
- Çyfeku, J. (2014):** *Introducing corpora in English language teaching classroom*. Paper presented at the Third International Conference on Education: Education Across Borders, held at the Faculty of Education and Philology, 'Fan S. Noli' University of Korce, Albania, published at *Vjetari Shkencor FEF, ISSN: 2078-7111*.
- Granger, S. (1996):** *From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora*. In K. Aijmer, B. Altenberg, & M. Johansson (Eds.), *Languages in contrast* (Lund Studies in English 88, pp. 37 – 51). Lund, Sweden: Lund University Press.
- Harrington, K. (2008):** *Perpetuating Difference? Corpus Linguistics and the Gendering of the Reported Dialogue*. In L. Litosseliti, H. Sauton, K. Harrington, & J. Sunderland (Eds.), *Theoretical and methodological approaches to gender and language study* (pp. 85-103). London/ New York: Palgrave Macmillan.
- Hunston S. (2002):** *Corpora in Applied Linguistics*. New York: Cambridge University Press: P. 2.

- Johns, T. (1991):** From printout to handout: Grammar and vocabulary teaching in the context of data driven learning. In T. Johns & P. King (Eds.), *Classroom concordancing* (Special issue of *ELR Journal* 4, University of Birmingham: Centre for English Language Studies), 27-45.
- Long, M. H., & Robinson, P. (1998):** Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second Language acquisition* (pp. 15-42). Cambridge: Cambridge University Press.
- O’Keeffe, A., McCarthy, M., & Carter, R. (2007):** From corpus to classroom: Language use and language teaching. Cambridge: Cambridge University Press.
- Pearson, J. (1996):** Teaching terminology using electronic resources. In S. Botley, J. Glass, T. McEnery, & A. Wilson (Eds.), *Proceedings of teaching and language corpora*. UCREL Technical Papers, (9), 22-39.
- Rayson, P., Leech, G and Hodges, M. (1997):** Social differentiation in the use of English vocabulary: Some analyses of the conversational component of the British National Corpus. *International Journal of Corpus Linguistics* 2 (1): 133-52.
- Reppen, R. (2010):** Using Corpora in the Language Classroom. In J. C. Richards, (Eds.) Cambridge: Cambridge University Press.
- Schmid, H.-J. and Fauth, J. (2003):** Women’s and men’s style: Fact or fiction? New grammatical evidence. Paper delivered at *Corpus Linguistics Conference*. UK: Lancaster University.
- Sheils, J. (1993):** Communication in modern languages classroom. Strasbourg: Council of Europe, Publishing and Documentation Service.
- Tribble, C. (2002):** Corpora and corpus analysis: New windows on academic writing. In J. Flowerdew (Ed.), *Academic discourse*. London: Pearson Education.
- Tribble, C., & Jones, G. (1997):** Concordances in the classroom: A resource guide for teachers. Houston: Athelstan Publications.
- Widdowson, H. G. (1990):** Aspects of language teaching. Oxford: Oxford University Press.
- Widdowson, H. G. (1998):** ‘Context, community and authentic language’. *TESOL Quarterly* 32/4: 705-16.
- Widdowson, H. G. (2000):** On the limitations of linguistics applied. *Applied Linguistics*, 21(2), 3-25.