



A Practical Corpus-based Approach to Teaching English for Tourism

Nuria Edo Marzá (Corresponding author)

Departamento de Estudios Ingleses, Universitat Jaume I

Avenida Sos Baynat s/n, 12071, Castellón, Spain

Tel: 0034 964 72 96 12 E-mail: nedo@ang.uji.es

Received: 10-09-2013

Accepted: 02-11-2013

Published: 01-01-2014

doi:10.7575/aiac.ijalel.v.3n.1p.129

URL: <http://dx.doi.org/10.7575/aiac.ijalel.v.3n.1p.129>

Publication framed within the Research Project awarded by the Ministerio de Economía y Competitividad (former Ministerio de Ciencia e Innovación), Reference FFI2011-2472, “Análisis léxico y discursivo de corpus paralelos y comparables (español, inglés y francés) de páginas electrónicas de promoción turística”. 2011-2014.

Abstract

The benefits derived from the use of corpora in the English for Specific Purposes (ESP) class seem out of question nowadays despite being a challenging task both for teachers and learners. Learning through a corpus-based approach or “data-driven learning” (DDL) (Johns 1986, 1994) allows students to get used to hands-on learning, thus promoting an inquiry-based approach leading them to discover patterns in authentic language use and making them linguistic researchers (Johns 2002) able to solve present and future linguistic needs of different nature with ease and autonomy.

Accordingly, the main aim of the research presented is twofold: firstly, to show the students of the ESP subject “Introduction to English for Tourism” the multiple applications and benefits that the use of a specialised corpus may have in the learning process of a foreign specialised language. With this aim, the 25 students who took part in the pilot study reported were presented the COMETVAL¹ corpus (Corpus Multilingüe de Turismo de la Universidad de Valencia), a specialised multilingual and multi-genre data base on tourism. In this initial stage, a workshop with different corpus activities was prepared to make students familiar with COMETVAL’s functioning, functionalities, and possible future professional applications.

Secondly, students responded a questionnaire designed *ad hoc* to know their perceptions about the use of specialised corpora in class –exemplified with COMETVAL’s workshop– and about their potential particular interest in devising, compiling and exploiting a similar specialised corpus in and for the class both for teaching/learning and future professional purposes.

Results show that students seem to feel positively motivated towards the use of corpora in class, specifically by the fact of having direct access to real, non-biased language samples from different genres; they do also find frequency and collocation-related data particularly interesting as well as linguistic patterns comparison between languages. In addition, students acknowledge the potential of autonomously relying on a specialised corpus for future professional queries.

Keywords: corpus, data-driven learning (DDL), English for tourism, English for Specific Purposes (ESP), COMETVAL

1. Introduction

1.1 Theoretical framework

According to Breyer (2008), teaching is a field in which technology has not had a strong impact, especially if compared with other areas. However, in the last 30 years the use of corpora in the classroom –understood as collections of naturally-occurring language texts, chosen to characterise a state or variety of a language (Sinclair, 1991) – has started to develop and research has shown that it can help improve students’ language proficiency (Varley, 2008). In fact, corpus benefits have actually been proven in studies on the teaching and learning of vocabulary by Cobb (1997), Cresswell (2007) and Stevens (1991). For instance, in the specific case of concordancing, it has not only proven to be a useful way “to mimic the effects of natural contextual learning” (Cobb, 1997, p. 314), researchers have also highlighted its usefulness for error correction in foreign or second language writing (Bernardini 2004; Chambers 2005; Gaskell & Cobb, 2004; Gray, 2005).

In the same way, the benefits derived from the use of corpora in the English for Specific Purposes (ESP) class seem out of question nowadays. In Belcher’s words (2006, p. 142), corpus-based analyses “pushes students toward empirically based understanding of language used for specific purposes” at the same time that they nicely complement existing reference works and may provide information which a dictionary or grammar book may not provide. In fact, given the specificity of ESP, many available materials in the form of textbooks are not easily adaptable to real learners’ needs

because of being too broad (Harwood, 2005). In addition, this kind of pre-elaborated material supposed to fit everybody's needs tends to be too artificial and lacks a final communicative objective and a consideration of students' specific features and necessities. The use of corpora in class as a regular -but not exclusive- practice constitutes one of the best and most feasible alternatives for surpassing these problems and for making the most of ESP teaching and learning:

The corpus-based approach [...] provides a means of handling large amounts of language and keeping track of many contextual factors at the same time. It therefore has opened the way to a multitude of new investigations of language use (Biber, Conrad and Reppen, 1998, p. 3).

Despite being a challenging task both for teachers and learners, using authentic and real-life examples with L2 learners is much more beneficial than using examples made up by the teacher, which do not foster real-life use of language (Madalena, 2001). In fact, McEnery and Wilson (1996) argue that foreign language teachers tend to produce or use simplified examples which do not reflect real-life use and which will, consequently, raise difficulties for students when confronted with real, more complex language that sometimes they are simply incapable of processing. In addition, learning through a corpus-based approach or "data-driven learning" (DDL) (Johns, 1986, 1994) makes it necessary and allows teachers to create classroom materials and activities to address specific class needs. In the same way, DDL allows students to get used to hands-on learning with authentic materials, thus promoting an inquiry-based approach leading them to empirically discover patterns in authentic language use and confronting the learner as directly as possible with the data, making him/her a linguistic researcher (Johns, 2002). In this way, students acquire the skills and are provided with the tools to be able to solve present and future linguistic needs of different nature with ease and autonomy.

Making students aware of the potential and mechanics of corpus use is the necessary step prior to teaching and learning through a corpus-based approach. In this way they will experiment and better understand, for instance, the importance of "appropriate lexico-grammatical patternings with consideration of various contextual and situational features of the discourse for the notions and functions one wishes to convey" (Flowerdew, 2008, p. 133). Moreover, students will get to know that corpora are invaluable resources when it comes to raise their awareness of language and textual patterning of genres (Tribble, 2002). Nonetheless, despite its many advantages and benefits, students do also need to be aware of the possible limitations or shortcomings of corpus linguistics since, as Bennett (2010) states, corpus linguistics is not able to provide all possible language at one time –no matter how planned or large a corpus is, it cannot be representative of all language–, it is not able to provide negative evidence, so that it can only tell us what is present in the corpus, and it is not able to explain why something is the way it is, it simply tells us what it is.

1.2 Objectives

The main aim of the paper here presented is twofold: firstly, to present to the students of the subject "Introduction to English for Tourism" from the Bachelor's degree in Tourism at Universitat Jaume I in Castellón (Spain) the multiple applications and benefits that the use of a specialised corpus exploited with a suitable concordancing program may have in the learning process of a foreign specialised language as well as its many applications from a future professional perspective. With this aim in mind students were presented the COMETVAL (note1) project and corpus and worked with a representative subsection of it (specifically of its English and Spanish subcorpora) through a series of pre-elaborated, "hands-on-corpus" tasks.

Secondly it was the researcher's objective to know students' perceptions about the use of specialised corpora in class derived from their own, first-hand experience previously acquired. With this aim in mind, students responded to a questionnaire on corpus use for teaching and learning purposes, that is, about its usefulness, difficulties, advantages and shortcomings as perceived by students aimed at determining the feasibility of compiling and using a specialised corpus in class designed and created *ad hoc* for the subject by the students.

Accordingly, the main aim of this study was to make students aware of and familiar with alternative and productive ways of learning a foreign specialised language, emphasising the need to "learn real language use from real language use" in a practical and empirical manner which fixes knowledge in a more durable and meaningful way.

1.3 The Multilingual corpus on Tourism of Valencia: COMETVAL

Students were thus introduced to corpus linguistics by showing them firstly what a corpus is and why it might be useful. This was done by presenting in class the corpus COMETVAL (Corpus Multilingüe de Turismo de la Universidad de Valencia), framed within an ambitious project aimed at getting to know better the discourse of tourism in order to strengthen the sector from a linguistic perspective. COMETVAL is stored in a specialised multilingual (Spanish, French and English), and multi-genre data base whose design and content allows its exploitation for terminological, discourse and metadiscourse analysis and translation purposes among others. COMETVAL includes institutional and private web pages related to the promotion of the industry of tourism. However, in this study we have focused on the independent subsection of COMETVAL devoted to the macrogenre "hotel webpage" (according to Calvi's (2010)) categorisation for considering it the most useful one for our students.

This big subsection of COMETVAL devoted to hotel web pages has a total of 7,577,102 words distributed into the different varieties of languages of work: American English, Irish English, UK English, American Spanish, Spanish from Spain, France French and Canadian French. The representative subsections of the corpus presented to students to work

with corresponds to a 750,000-word sample from the UK English part of the corpus and to a 750,000-word Spanish (from Spain) subcorpus employed for comparison purposes.

Students were also shown how COMETVAL's data base includes many fields and subfields allowing the user to filter very specific information. Among the fields, we find those related to administrative organization and identification such as those of title of the document, link, ID number of the register and researcher introducing the register. The source, location and language are also included as well as the genre and macrogenre each sample belongs to. Apart from this, the text itself together with the specific field and subfields dealt with in it are also included as well as the category of the establishment and the kind of addressee, among others.

2. Method

2.1 Participants and instruments

The research was conducted in two two-hour sessions (four hours in total) in a class with 25 students with a presumed B1 level (according to the Common European Framework of Reference for Languages) and 6 computers available with the concordance program WordSmith Tools 5.0 (WST) (note 2), the one chosen for corpus exploitation and analysis.

2.2 Procedure

The procedure followed comprehended the following stages:

Firstly, in an initial session, the COMETVAL project was presented to students by showing them the data base in which it is stored, how it works, its functionalities, the languages it works with and the different fields and subfields contained in each register. In fact, they were explained how COMETVAL allows users to filter the kind of data they want to work with. They were shown how, for instance, if they were interested in retrieving texts dealing with hotel facilities in high category hotels with the variety American English they would have had to do a search according to the parameters American E + Facilities (*instalaciones*) + high category hotel (*hotelero alto*). The search was carried out in class in order to illustrate the explanation.

Secondly, students were explained that in order to work with a concordance program (in this case the aforementioned WST 5.0), COMETVAL's texts need to be exported to a plain text format. Since COMETVAL is a corpus belonging to a research group and therefore not open access, I retrieved only a part of it for students to work with in class: a 750,000-word UK English subcorpus plus a 750,000-word Spanish (from Spain) subcorpus (for comparison purposes). Therefore, students did not work with the whole COMETVAL but with a representative and big enough sample of it.

Thirdly, students were arranged in 5 groups of 4 and a group of 5 in the 6 computers available with a WST 5.0 licence. Previously, students had been asked to read a simple tutorial about the basic use of WST. At the end of the session a brief hands-on, warming-up practice –in which results interpretation was also explained– was carried out with WST, more specifically with its application Concord, the one our activities were focused on.

In the fourth stage, during the second session, students, who had a folder with the corpora to work with in the desktop of their computers, started to work with a series of corpus familiarisation and exploitation tasks, namely 5 brief tasks (shown in the results section) with which to explore the potential of corpus usage both for teaching and learning purposes and for prospective professional use.

Finally, after the corpus familiarisation and exploitation tasks, students were asked about their impressions on corpora use –its usefulness, perceived difficulty, etc.– and about their views on the possibility of compiling and exploiting corpora in class as a regular practice. The questionnaires designed for such purpose and the results obtained are also presented in section 3.

3. Results and discussion

3.1 Corpus familiarisation and exploitation activities developed in class

The resulting activities designed and developed in order to familiarise students with corpus linguistics and in order to show them its usefulness for teaching and learning purposes are the ones below. A brief outline of the results obtained or to be obtained has also been included. Obviously, the corpus students worked with was a written one so that these activities are mainly (but not only) intended at improving the written competence of students.

Summary of activity 1:

The adjective *friendly* is becoming increasingly used and common among hoteliers and hotel web pages to designate that their hotels are showing kind interest and goodwill towards specific types of clients. Get familiar with the use of this term by answering questions 1 to 4:

1. How many concordance lines can you find with the term “friendly”? Do you think it is a relevant term in the field? Why? Why not?

282 concordance lines were retrieved with the term *friendly*. Students all agreed, due to its frequency and thus to its relatively common use in the corpus, that it was a relevant term in the field.

2. Obviously, you will find many R1 collocates for “friendly” in the corpus (of the “staff” (*friendly staff*) or “atmosphere” (*friendly atmosphere kind*). Focus now on L1 collocates (without dismissing R1 ones) and note down what other things or aspects can be “friendly”. Note down the number of instances of each L1 collocate in the corpus in order to identify the most frequent friendly-type collocations.

The results obtained are the ones shown in Table 1.

Table 1. L1 collocates for the search word “friendly”

L1	Search word	R1
Family (132)	friendly	Hotel, holiday park, resort, dinner, services, guesthouse, inn, value hotel, etc.
Pet (29)		Places, establishment, cottages, hotel, inn, guesthouse.
Dog (12)		Areas, rooms.
Gay (1)		Hotel.
Wheelchair (5)		Hotel, hotel access.
Cyclist (1)		Hotel.
User (1)		Booking.
Child (5)		Areas, menus, hotel.

Table 1 shows, for instance, resulting collocations of the “family(-)friendly hotel” or “pet(-)friendly establishment” type together with the frequencies of joint occurrence between the L1 collocate and the search word.

- Look at the concordance lines obtained and discuss in small groups the most suitable Spanish translation for the term “family friendly hotel”.

Among the options discussed by students, we find: *HOTEL apropiado/adequado/apto para familias, amigo de las familias, amigable para las familias, para toda la familia, para familias, hotel familiar.*

- Use the Spanish version of the corpus and look for the possible translations mentioned in 3. According to corpus results, are the translations proposed a reflection of real use?

Some of them were and some were not. According to students, frequency criteria indicate the preferred uses (in bold):

HOTEL familiar (56)

HOTEL para toda la familia (28)

HOTEL para familias (13)

HOTEL apropiado (5) /adequado (2) /apto(0) para familias,

HOTEL amigo de las familias (0)

HOTEL amigable para las familias(0)

Summary of activity 2:

- Brainstorming: Imagine you are writing an e-mail to a client, how would you ask a client to ask you whatever he/she may need?

These were some of the answers provided by students: *Do you have any question? If you have a question, tell me. I can answer your questions. I can help you with your questions.*

- Look for as many synonyms for the word *question* as you can.

These were mostly the terms retrieved: *question, query, doubt, inquiry* and *enquiry*.

- Using Concord, look for the word *question* (and the synonyms found in 2) in the corpus. Do not forget to add an asterisk (*) at the end of the search word (for *query, inquiry* and *enquiry* use *quer**, *inquir** and *enquir**) so that the program retrieves both singular and plural forms. Note down all the expressions that in your opinion may be useful for you to know in order to ask clients about their questions or doubts:

Some of the concordance lines obtained and thus some of the responses given were the ones listed below. As previously mentioned, students were also encouraged to create a sort of manual of useful phrases for everyday use at work in different situations.

If you have any questions , our customer service staff will be more than happy to provide assistance.
Any questions regarding claims should be directed to customer.relations@thistle.co.uk
the friendly staff at Park Plaza County Hall is happy to answer any question and accommodate any request.
If the question you have is not listed below please do not hesitate to contact us,
If you have no additional questions about the offer(s), you can help us speed up the booking process by submitting credit card details in your reply e-mail
If you have any requests concerning your personal information or any queries with regard to the capture, storage and use of personal information please contact Marketing
For current membership queries please contact Katie Henry on: 020 7973 7550
if you have any queries relating to your booking or would like to contact us with any special requests, please complete the form
For the avoidance of doubt , prices do not include any incidental charges, which you may incur during your stay
Limited accommodations for guest pets (Please inquire)
We will then use that information to assist in any inquiries about your transaction.
Please complete booking enquiry form to check availability.
Please enquire for further information.

Please enquire for details about booking for over/under four people.
Email rghbrid@btinternet.com to enquire about availability.
Please contact Anna for all enquiries and for wedding pack information.
Our Concierge will be delighted to assist with any enquiries you may have including theatre or travel bookings.

Summary of activity 3:

1. In Spanish, the verbs *hospedar*, *acomodar* o *alojar* (“to accommodate” in English) have either a pronominal use or are followed by the Spanish preposition “a” (which in English tends to be “to”). Which of the two options below is the correct one?

The hotel accommodates to 150 guests this week. Option chosen by 16 students.

The hotel accommodates 150 guests this week. Option chosen by 9 students.

2. Check in the corpus the use of the verb “to accommodate” and note down some significant concordance lines that illustrate its use. Check whether your answer to 1 was correct.

3.

Our apartments accommodate 2- 8 people making us great for both couples, friends, family and small groups.
The Old Dispensary accommodates 2-4 adults in a double bedroom and on a double sofa bed.
We cannot accommodate guest pets

4. You have probably also come across the expression “accommodate up to” among the concordance lines retrieved. What is its meaning according to the corpus? Note down a couple of these examples for you to keep.

Students deducted from the examples in the corpus that “accommodate up to” means *alojar/acomodar/hospedar HASTA X personas/comensales*, etc.

The Ballroom can accommodate up to 500 guests for receptions.
The Dog and Partridge accommodates up to 16 people and is available to group bookings.

Summary of activity 4:

1. Look for the adjectives most frequently used to describe location by introducing the search word “location*” in WST.

In this case the students came across adjectives such as perfect, ideal, quiet, rural, lakesite, dazzling, peaceful, great, incredible, superb, prime, desirable, central, etc. This simple kind of activity allows them to enlarge their vocabulary knowledge and enrich their positive evaluative language, so important in the tourism industry.

Summary of activity 5:

1. The verb “to contact” (understood as *ponerse en contacto* (getting in touch with somebody through social interaction)) may cause some trouble among Spanish students. In your opinion, is the following sentence correct or incorrect? *Contact with me if you want more information.*

19 students believed it to be incorrect whereas 6 students considered it correct.

2. Look for the word “contact” in the corpus and note down some evidence of the correct use of the word. Note down as many patterns as you can with the word “contact”. Was your answer to question 1 correct?

Contact + D. O.
If you are an existing customer, we will only contact you by electronic means with information about services
Please contact us to check availability and to make reservations.
Please contact the Concierge for details.
Contact for:
Rezidor Hotels ApS Danmark may be contacted for additional information under 3.
Contact on:
If you would rather speak to us about a booking we can be contacted on +44 (0)1784 220961

3. Is there any instance in which “contact with” is correct?

Yes, with the structure “keep in contact with” (*mantenerse en contacto con alguien*).

So you can keep in contact with anywhere in the world from anywhere in the hotel.
--

3.2 The prospective compilation and use of a specialised corpus in class: the questionnaire on students' perceptions

In the final stage of the research students were administered and responded an anonymous questionnaire designed *ad hoc* to know about their perceptions on the use of specialised corpora in class and about their particular interest in devising, compiling, implementing and exploiting a similar specialised corpus both for teaching/learning and future professional purposes.

Table 2. Questionnaire administered to students and number of responses provided for each item.

PART 1 – Perceptions and opinions about COMETVAL and corpus familiarisation activities.	1 No	2 Dubious	3 Yes
1. Do you find COMETVAL useful for you as a student?	0	4	21
2. Could COMETVAL be useful for you as a future tourism professional?	2	4	19
3. Do you find it useful the way COMETVAL's data base has been designed (its layout)?	2	5	18
4. Do you find having access to corpora in different languages useful?	0	0	25
5. Do you find having access to different genres useful?	0	0	25
6. Do you find COMETVAL'S data base manageable?	1	9	15
7. Do you find concordance software useful?	1	4	20
8. Do you find concordance software manageable for students?	2	7	16
9. Would you be able to work with concordance software on your own?	2	8	15
10. Can a corpus help you to identify new and relevant uses of a word and their possible translation? (Task 1)	0	3	22
11. Can a corpus help you to identify and use set expressions in order to express yourself more naturally? (Task 2)	0	2	23
12. Can a corpus help you to find out the correct way of using a verb avoiding (incorrect) literal translations? (Task 3)	0	1	24
13. Can a corpus help you to make your discourse more adequate and vivid by providing, for example, instances of evaluative language (adjectives) in use? (Task 4)	0	2	23
14. Can a corpus help you to check whether a given use of language is correct and how to use it correctly? (Task 5)	0	1	24
PART 2 – Perceptions and opinions about the possibility of compiling and using <i>ad hoc</i> corpora in class on a regular basis.	1 No	2 Dubious	3 Yes
15. Would you find the use of corpora designed <i>ad hoc</i> useful for you as a student?	2	1	22
16. Would you find the use of corpora designed <i>ad hoc</i> useful for you as future tourism professional?	4	1	20
17. Would you like to participate in the compilation of a corpus for teaching/learning purposes?	2	8	15
18. Would you find it useful to use corpora in class on a regular basis?	3	0	22
19. Would you like to be taught only with a corpus?	15	4	6
20. Would you like to be taught combining textbook exercises, corpus tasks and other resources?	0	0	25
21. Can a corpus help you gain competence in speaking?	0	3	22
22. Can a corpus help you gain competence in writing?	0	2	23

As may be observed in Table 2, the questionnaire designed shows a very simple three-item scale for the 22 items posed. The questionnaire included two separate but complementary parts according to the objectives of the study. The resulting number of responses obtained for each of the items has been included in Table 2 to make results fully available and illustrative

Regarding the perceptions and opinions about COMETVAL and the concordance software familiarisation activities carried out in class, most students seemed to understand the scope and potential of corpus use since they acknowledge

the usefulness of COMETVAL for both present (learning) and future (professional) purposes (items 1 and 2). Most students do also see the benefits of storing corpora in multi-field registers allowing for quick and meaningful information retrieval –based on language, genre, etc.– so that they, once again, acknowledge the benefits of COMETVAL’s layout and retrieval possibilities (items 3 to 5). Nonetheless, 9 students are dubious about the manageability of COMETVAL (item 6). As regards concordance software, despite the fact that WST is not the most user-friendly of concordancers, 80% of the students find it useful, 64% find it manageable and 60% of them believe they would be able to work with it on their own (items 7 to 9). Finally, items 10 to 14 in the first part of the questionnaire tried to make students reflect about the purpose and usefulness of the familiarisation activities carried out in class. The vast majority of students agreed that a corpus is a useful tool in order to identify new and relevant uses of a word and their possible translation, identify and use set expressions in order to express themselves more naturally, find out the correct way of using a given word (thus avoiding incorrect literal translations) and make their discourse more adequate and vivid by providing instances of, in this specific case, evaluative language (adjectives) in use.

As regards the second part of the questionnaire devoted to perceptions and opinions about the possibility of compiling and using *ad hoc* corpora in class on a regular basis, the majority of students find the use of corpora useful for them both as students and as prospective tourism professionals (items 15 and 16). 60% of them would like to participate in the compilation of a corpus for teaching/learning purposes (item 17). 88% of the students would you find it useful to use corpora in class on a regular basis but only 24% of them would like to be taught exclusively with a corpus (items 18 and 19). Nonetheless, 100% of the students agreed that they would like to be taught combining textbook exercises, corpus tasks and other resources and the vast majority also agreed that a corpus could help them gain competence in both speaking and writing (items 20 to 22).

Accordingly, to sum up, results show that, in general, students feel positively motivated towards the use of corpora for learning purposes. They specially like the fact of having direct access to real, non-biased language samples from different genres; they feel motivated by the many and different searches that can be undertaken due to the inclusion of retrieval categories and they do also find particularly interesting collocational results as well as the data derived from the comparison between English and Spanish patterns. In addition, students acknowledge the potential and benefits of autonomously relying on a specialised corpus for future professional queries and seem positively motivated towards this possibility even though they prefer not to be taught exclusively with a corpus but with a combination of methods. Students seem positively motivated towards compiling and using a corpus in class because a great majority of them do believe it could help them to improve their linguistic competence, specifically as regards their oral and written skills. Results show then a general positive attitude towards the study and, hence, towards the use of a corpus-based approach for ESP teaching and learning, which could be translated into a combination of traditional and corpus-based methods.

4. Conclusion

As suitably formulated by Aston (2000, p. 16) “language pedagogy is increasingly designing its own corpora to its own criteria”. We do not know exactly how these criteria will develop in the next few decades. One thing that we can be sure of, however, is that the field of corpus linguistics and language teaching has an exciting future that both researchers and teachers can, and should, look forward to. ESP is probably one of the most characteristic applications of corpus linguistics, with or through which the areas of register, lexicogrammar and phraseology can be analysed also in specific discourse.

Corpus linguistics and terminotic tools offer the possibility of introducing students to a hands-on, task-based approach to ESP –in the case of this study, the language of tourism– that makes them more aware of real language use and makes them more autonomous and as well as deductive and even curious. As has tried to be shown throughout this study, the possibilities offered by corpus analyses are varied and complementary, these including the possibility of verifying certain hypotheses about languages, finding support examples or counter examples, using a corpus as a reference, seeing whether certain instances are possible or which one is more natural, answering specific questions of certain words, phrases and structures, knowing where the problems are or what should be taught, extracting generalisations and comparing languages among many others

Learning through a corpus allows students to get involved in a “self-discovery” process that enhances active and lasting learning. They have the possibility of analysing the language *in vivo* and not “canned” in aseptic and artificial text books and this is something students appreciate if we help them to understand the mechanics of corpus exploitation and analysis.

Obviously, learning through a corpus-based approach implies a necessary ICT (Information and Communication Technologies) component translated into the need of using computers and software concordance programs which are not always available or are not always feasible. However, if the material means needed are at students and teachers’ disposal, data-driven learning is probably one of the most effective, meaningful and long-lasting ways of learning and students seem to perceive it in the same way. In the researcher’s opinion and according to the results obtained all through this research, if students know “how” to retrieve and interpret data from a corpus, they will easily draw conclusions, relate data and make generalisations as regards language use, something which will make them more autonomous and reflective in their way of approaching any foreign language.

References

- Aston, G. (2000). Corpora and language teaching. In L. Burnard & T. McEnery (Eds.), *Rethinking Language Pedagogy from a Corpus Perspective* (pp. 7-17). Frankfurt: Lang.

- Belcher, D. (2006). English for specific purposes: Teaching to perceived needs and imagined futures in worlds of work, study, and everyday life. *TESOL Quarterly*, 40, 1, 133-156.
- Bennett, G. R. (2010). *Using corpora in the Language Learning Classroom: Corpus Linguistics for Teachers*. Michigan ELT.
- Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. McH. Sinclair (Ed.), *How to use corpora in language teaching* (pp.15-36). Amsterdam: Benjamins.
- Biber, D., Conrad, S & Reppen, R. (1998). *Corpus linguistics: investigating language structure and use*. Cambridge: Cambridge University Press.
- Breyer, Y. (2008). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153-172.
- Calvi, M.V. (2010). Los géneros discursivos en la lengua del turismo: una propuesta de clasificación. *Ibérica*, 19: 9-32.
- Cobb, T. (1997). Is there Any Measurable Learning from Hands-on Concordancing? *System*, 25(3), (pp. 301-315).
- Cresswell, A. (2007). Getting to 'Know' Connectors? Evaluating Data-driven Learning in a Writing Skills Course. In E. Hidalgo, L. Quereda and J. Santana (eds.), *Corpora in the Foreign Language Classroom* (pp. 267-288). Amsterdam: Rodopi.
- Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning and Technology*, 9 (2), 111-125.
- Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System*, 32, 301-319.
- Gray, B. E. (2005). Error-specific concordancing for intermediate ESL/EFL writers. Retrieved January 18, 2013 from <http://dana.ucc.na4.edu/~bde6/coursework/Projects/PedagogicalTip/homepage.html>.
- Harwood, N. 2005. What do we want EAP teaching materials for? *Journal of English for Academic Purposes*, 4: 149-161.
- Johns, T. (1986). Microconcord: A Language-learner's Research Tool. *System* 14(2), 151_162.
- Johns, T. (1994). From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-driven Learning. In T. Odlan (Ed.), *Perspectives on Pedagogical Grammar* (pp. 27-45). Cambridge: Cambridge University Press.
- Johns, T. (2002). Data-driven Learning: The Perpetual Challenge. In B. Kettemann & G. Marko, *Teaching and learning by doing corpus analysis* (pp. 107-117). Proceedings of the Fourth International Conference on Teaching and Language Corpora, Graz 19-24 July, 2000.
- Flowerdew, L. (2008) *Corpus-Based Analyses of the Problem-Solution Pattern. A phraseological Approach*. Amsterdam/Philadelphia, John Benjamins Publishing Company.
- McEnery, T. & Wilson, A. (1996). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Maddalena, S.R. (2001). *An investigation into how corpus analysis may be used in the second language classroom to solve some of the problems surrounding non-native speakers' understanding of seemingly synonymous words*. Retrieved August 15, 2013 from ERIC database.
- Stevens, V. (1991). Concordance-based Vocabulary Exercises: A Viable Alternative to Gapfillers. In J., Tim & P. King (Eds.), *Classroom Concordancing* (ELR Journal 4), (pp. 47-63).
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Tribble, C. (2002). Corpora and corpus analysis: New windows on academic writing. In J. Flowerdew (Ed.) *Academic Discourse* (pp. 131-149). Harlow, England, Longman.
- Varley, S. (2008). I'll just look that up in the concordance: Integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning*, 22 (2), 133-152.

Notes

Note 1. COMETVAL is the acronym for Corpus Multilingüe en Turismo de Valencia.

Note 2. The author is aware of the fact that WordSmith Tools is not probably the most user-friendly concordance program currently available but believes it to be the one with the greatest potential and this is the reason for having used it.