

## ADVANTAGES OF USING CORPORA TO TEACH ENGLISH

**Alcántar Díaz Carlota de Jesús**  
Alcantardiaz2005@hotmail.com

**José Luis Quintero Carrillo**  
jlquinterocarrillo@gmail.com

**Universidad Autónoma de Nayarit**

### **Abstract**

Language teaching in the last three decades has focused on the process of learning through communication, with little or no attention to the forms of the language. However valuable this approach, there has been growing recognition that knowledge of the language system and accuracy are also relevant. Teachers' interest in the content of language learning has revived and it is precisely corpora-based research that makes this information available for language teachers and researchers (Kennedy, 1998). This article provides some examples of corpora-based works and mentions some advantages of using corpora in the English learning classroom, it reviews several studies that have been carried out on this topic (Swales, 1998, 2002; Watson Todd, 2001; Schmitt, 2002; Gabrielatos, 2005) that explain why using corpora for English Teaching is valuable.

**Key words:** corpora, advantages, English, teaching.

### **Introduction:**

Traditional English teaching (emphasizing systemic accuracy) has been all but abandoned for the last three decades and new theories have emerged; language teaching through communication has become popular. The communicative approach to teach English highlights the importance of communicating ideas without worrying about the way they are expressed, that is to say, the message is more important than systemic language accuracy (Kennedy, 1998). Due to the shift in teaching focus, the teachers' role changed as well; in the traditional method teachers were the source of information, while in the communicative approach they became guides, organizers, and knowledge facilitators.

In spite of the value given to English learning through communication (focused especially on oral fluency), teachers and researchers nowadays recognize that systemic language accuracy is relevant, therefore the teachers' interest in the content of language learning has revived, and it is precisely the access to Corpora (language data bases) that informs teachers of authentic use of language and provides them with valuable examples for their teaching practice (Kennedy 1998).

The information about the distribution of language elements and linguistic process can influence pedagogy in different ways (Kennedy, 1998); the analysis of a corpus (a single data base) can influence the content of a language class since teachers can give priority to certain contents that are more frequent in a language system, corpus analysis is a guide for selection of language elements and the importance each of these elements should be given in a language class.

Corpora studies could also influence English teachers' methodology; teachers could plan activities in the lesson for the students to explore the language using computer programs that allow them to make their own conclusions about the language. An example of these programs is the *concordancer* that is used to identify words in a particular corpus and their usual collocations.

In general, we can say that technology use has the potential to radically change the way teachers plan and conduct a language class. Some of the pedagogical implications of Corpora are that teachers can use data bases to inform and plan their classes and students can participate in their own language learning by exploring a corpus and discovering language use in real life.

### **Advantages of using corpora.**

Corpora analysis for language teaching and the role it plays in the teaching-learning process has both advantages and disadvantages. This article however is limited to showing some of the advantages of using corpora to inform linguistic practices in the English classroom. It should be noted that the advantages mentioned below are just a small sample of a larger list.

#### **a) Corpora can inform deductive and inductive approaches to English teaching:**

In the deductive approach to teaching English, corpora analysis provides evidence that informs teachers (especially to those who are non native speakers) about the use of language elements they are presenting in class and provides them with clear and authentic examples of the language elements.

In the inductive approach to teaching English, corpora analysis provides students with data to infer language rules by themselves.

#### **b) The use of corpora improves English programs and materials design courses**

Gabrielatos (2005) suggests that teachers and material writers may unwittingly present their personal informal observations about language as the true and full picture of language structure and use, or present their own preferred usage as the only 'correct' or 'acceptable' one' (p.5).

However, because corpora informs us about the use of language elements in a way that our intuition cannot, corpora can enlighten the syllabus, and most teaching materials can be based around corpus data. A clear example of this is the use of modals and the form in which they are presented in English textbooks. According to the corpora-based studies on modals (e.g. *would, can, might* etc.) carried out by several researchers (Holmes 1988; Hyland 1994; McEnery & Kifle 2002; Römer 2004) in which they compared how speakers and writers use this language to how textbooks claim students should use it; they found that textbooks are not teaching the full inventory of modal language, they are also providing confusing explanations for some of the language they teach.

Hyland (1994) concludes that for the most part, modal expressions are simply introduced without system or comment and are summarily dealt with in a single exercise which fails to emphasise either their function or importance. Generally, the range of modal verbs addressed and the information provided on their use is inadequate.... (p. 247).

In the same vein, Williams (1988) affirms that the selection of examples is unclear, but she would suspect that authors too enthusiastically use introspection or a type of educated hunch, instead of an empirical research.

Another example related to English teaching material design and its flaws is the honest review Swales wrote about his book *Academic writing for Graduate Students* (Swales, and Feak, 1994), in which he recognizes the faults with the textbook syllabus. Swales and his co-author wrote the book before linguistic computerized analysis became a tool for linguists and material designers. They affirm that the book was written “before we became aware of corpora, and was based on our own lengthy experiences as writing instructors” (Swales 2002, p. 152). Since Swales had no corpus data to inform his material, students are given a confusing picture of how imperatives (e.g. *see, notice, suppose, consider*) are used in academic writing. Swales et al. (1998) demonstrated this by comparing the imperatives in his textbook with the more common imperatives found in a corpus of academic writing that contained texts from different areas: art history, chemical engineering, communication studies, experimental geology, history, linguistics, literary criticism, philosophy, political science, and statistics.

After the corpus analysis he concludes that from ten lexical choices; *notice, imagine, refer, observe, take the case of,* and *disregard* either did not occur in his corpus or occurred no more than twice in the main text. Swales et al. (1998) observe that verbs like *suppose* may occur in mathematical arguments, but with more frequency in non-mathematical philosophy. They do not mention *see*, and speculate that *consider* is “probably rare outside (philosophical) arguments, whereas the current data suggest its common use in at least the major school of theoretical linguistics.” (Swales et al 1998, p. 118).

### c) Corpora and language awareness

Language awareness can be defined as “the development in learners of an enhanced consciousness of and sensitivity to the forms and functions of the language” (Carter 2003, p. 64). This is why corpora analysis plays an important role in learners language awareness: when students work with corpora they discover language for themselves. Bolitho et al. (2003) state that language discovery is the key element of the language awareness approaches. Van Lier (2001) gives an example of a language awareness teaching activity which shows how the focus is on the learners, not the teacher:

‘...using data provided or collected, learners observe and analyse patterns of interest and come up with descriptions or tentative rules, usually in group work. In most cases the data are from authentic sources.... [...] Teachers can also use concordancers with authentic texts in order to raise awareness of grammatical, stylistic and lexical features....’ (Van Lier, 2001, p.164)

In language learning, reflection and awareness-raising could be associated with 'noticing'. An important function of raising language awareness in the learners is to help them 'notice' the language feature' and 'notice the gap' between their own production and the correct grammatical feature as produced by native speakers (Schmidt, 1990). An example of 'noticing' is a class activity Watson Todd (2001) carried out, he identified vocabulary that his Thai students in a pre-intermediate course had problems with in the writing practices, but he didn't correct these problems, just underlined them. He asked learners to search for this vocabulary on the Internet, and to make concordances (10 examples) containing the vocabulary. Then he asked them to find out why they made mistakes by comparing their writing and the concordances sentences, after looking at concordances of the word *capable* a student wrote the next correct rule:

✓ *Capable* is used between verb *to be* and *of*

✓ *Capable* is always followed by *verb + ing*

This student also corrected his incorrect use of the word *capable* successfully. The results of this English practice was that at the end of the lesson learners were able to notice and self-correct their errors in **over 78%** of cases.

#### **d) Corpora and ELT methodology.**

Corpora is normally associated with noticing and language awareness. Gabriëlatos (2005) draws attention to the fact that corpora can also be used by teachers following more traditional methodologies, such as PPP (Presentation—Practice—Production), as well as task-based approaches. He argues this because corpora can be used to look up examples of the language the teacher (or the students) wish to focus on, for instance in a PPP lesson, instead of using invented sentences which contain the target language (e.g. *should* and *must*), the teacher could obtain sample sentences containing *should* and *must* from a corpus, then teach the lesson in the typical fashion.

However we should not forget that although corpora can be used to design a more traditional style of lesson like PPP, corpora use is normally seen by researchers as helping to create a less teacher-centred classroom atmosphere.

#### **e) Corpora develops students and teachers autonomy and independence**

Different investigations suggest that using corpora promotes students' autonomy, because students can take responsibility for their own language learning, looking for language features in the different existing language data bases with the teachers' guidance. Instead of relying on the teacher for information, corpora can be used by learners to find things out for themselves.

Students' autonomy is not a threat to teachers. Gabriëlatos (2005) argues that teachers' roles in corpus-based classrooms won't be any less important: their role will only be different

This is not to say that the teacher's role is diminished; rather, it is enriched and diversified. The teacher becomes less a provider of input and facts about language and more a facilitator and consultant, or, at the learner-centred end, a co-researcher. (Gabriëlatos, 2005, p.19).

Corpora use also empowers non-native teachers because it helps them to be independent. For instance they can find suitable examples of 'real life' language which makes them feel more confident about the language they are presenting to the students.

#### **Conclusions.**

Nowadays in a globalized world it is impossible to conceive of education without technology and all the advantages that it has given to the teaching- learning process in general, and to the language teaching-learning process in particular. It is difficult not to realize that computers are used to inform teachers in developing classroom activities as well as to facilitate their job.

The use of corpora is a new tool that provides teachers with authentic data about language structure and also promotes student autonomy because they can explore a determined corpus and do their own research about language features.

Unfortunately, using corpora is not an easy task, especially for those teachers from institutions whose aim is not just English teaching (typical public school with limitations), and therefore therefore it is possible that even when English teachers are informed of this new technology, they may continue to base their teaching on textbooks with dubious language content.

Schools require investing in computer equipment as well as in software to analyse corpora in order to introduce this new technology and teachers and students can make use of its advantages.

Despite the fact economical problems (lack of computers in the school) would be enough reason for teachers not to use corpora to supplement their teaching course; they can use corpora in a basic way at home with the help of one computer and internet using the free downloadable software to analyse a corpus; they can identify examples of authentic language use for their regular classes no matter the methodology they use to teach.

## References

- Bolitho, R, Carter R, Hugues R, Ivanic R, Masuhara H, Tomlinson B. (2003) Ten questions about language awareness. *ELT Journal* 57(3): 251-259.
- Carter, R (2003) Language awareness. *ELT Journal* 57(1): 64-65.
- Kennedy, G. (1998) *An introduction to corpus linguistics*. Longman.
- Gabrielatos, C. (2005) Corpora and language teaching: just a fling or wedding bells? *TESL EJ* 8(4): A1. Retrieved from [http://scholar.google.com.mx/scholar\\_url?hl=es&q=http://eprints.lanacs.ac.uk/530/1/TESL-EJ\\_8\(4\)-A1.mht&sa=X&scisig=AAGBfm3wQWjnh\\_8MVXf7AVML0P6\\_a2zow&oi=scholar&ei=WC0yU6S5LOri2AXewoH4Ag&ved=0CCoQgAMoADAA](http://scholar.google.com.mx/scholar_url?hl=es&q=http://eprints.lanacs.ac.uk/530/1/TESL-EJ_8(4)-A1.mht&sa=X&scisig=AAGBfm3wQWjnh_8MVXf7AVML0P6_a2zow&oi=scholar&ei=WC0yU6S5LOri2AXewoH4Ag&ved=0CCoQgAMoADAA)
- Holmes, J. (1988) Doubt and certainty in ESL textbooks. *Applied Linguistics* 9(1): 21-44.
- Hyland, K. (1994) Hedging in academic writing and EAP coursebooks. *English for Specific Purposes* 13(3): 239-256.
- McEnery, T. y Kifle, N.A. (2002) Epistemic modality in argumentative essays of second-language writers. In J. Flowerdew (ed.), *Academic Discourse*. Harlow: Longman, pp.182-195.
- Römer, U. (2004) A corpus-driven approach to modal auxiliaries and their didactics. In J. Sinclair (ed.), *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins: 185-199.
- Schmidt, R. The Role of Consciousness in Second Language Learning. *Applied Linguistics* 11(2), 129-158 (1990)
- Schmitt, N. (2002). Using corpora to teach and assess vocabulary in Tan, M. (ed.), *Corpus Studies in Language Education*. IELP Press.
- Swales J. M. y Feak C.B. (1994) *Academic Writing for Graduate Students*. Michigan Series in English for Academic & Professional Purposes
- Swales J. M., Ahmad U., Chang, Y-Y, Chavez, D., Dressen D., y Seymour R. (1998) Consider this: the role of imperatives in scholarly writing. *Applied Linguistics* 19(1) 97-121.
- Swales J.M (2002) Integrated and fragmented worlds: EAP materials and corpus linguistics. In J. Flowerdew (ed.) *Academic Discourse*. Harlow: Longman:150-164.
- Van Lier L. (2001) Language awareness. In R. Carter & D. Nunan (eds.), *The Cambridge Guide to teaching English to Speakers of Other Languages*. Cambridge: Cambridge University Press:160-165.

- Watson Todd R. (2001) Induction from self-selected concordances and self-correction. *System* 29: 91-102.
- Williams M. (1988) Language taught for meetings and language used in meetings: is there anything in common? *Applied Linguistics* 9(1): 45-58.