

THE ROLE OF ENGLISH CORPUS IN THE FIELD OF EDUCATION

Yusupova Ra'no Davronovna

Buxoro Davlat Universiteti, xorijiy tillar fakulteti,
ingliz adabiyotshunosligi yo'nalish talabasi.

Abstract:

The article describes the role of the English corpus in the field of education and processes in the field of international education.

Keywords: ELT, documented, Corpus, linguistic, elements, independence.

Helping learners, teachers and teacher educators improve their awareness of how language works is a key goal in the field of English Language Teaching (ELT) worldwide. From everyday conversations, to business meetings and academic lectures, access to all kinds of digital language data is now possible using a corpus. A corpus is a digital collection of textual examples created for the purposes of exploring how language is used (Sinclair, 2004). The millions of words and phrases in a corpus can be searched and retrieved instantly, using similar technical skills as those required by standard search engines. The benefits of using corpora in language teaching and learning have been well documented in the literature such as the provision of better quality language input for learners; the development of more accurate grammatical and vocabulary knowledge; and the design of more reliable teaching materials; the facilitation of students' and teachers' control over their own learning and teaching. Surprisingly, the potential of corpora in ELT remains largely untapped.

Spoken and written corpora

Corpus linguistics provides us with many tools for the study of language, its structure and patterns. Even a cursory analysis of the field reveals that there exist many written and spoken corpora (O'Keeffe, McCarthy and Carter 2007). One of the most widely cited corpora is the British National Corpus (BNC) that consists of 100 million words (of both written and spoken English). The written data constitute ninety per cent of the whole corpus (newspapers, books, letters, essays) and spoken data (business meetings, phone-ins, radio shows) amount only to ten per cent. There are



also corpora of American English. The Corpus of Contemporary American English (COCA) is a freely available corpus that contains 410 plus million words. It was developed between 1990 and 2010 by adding 20 million words each year from both spoken and written English (spoken, fiction, magazines, newspapers, academic texts). Compiling a spoken corpus, in comparison with a written ones, is a much more difficult task requiring recording data and carefully transcribing it. Due to this fact, as Aijmer (2002) reports, among the existing corpora written ones are more prevalent and spoken corpora are fairly small.

How can corpora be used in English Language Teaching (ELT)

The advent of corpora has provided new methodologies for language study and changed linguists approach in areas such as lexicography or English Language Teaching (ELT). With regard to ELT, there are several aspects of language pedagogy which demonstrate how corpus findings have been put to practical use. First of all, all major publishers use corpus data to compile dictionaries and teaching materials. If one wishes to publish a dictionary that reflects the way English is spoken every day, they need to obtain accurate information about the language and the only way to do so is to investigate the authentic use of English by its speakers. Corpora are a reliable source of authentic language data and therefore they serve as a basis for developing a wide range of pedagogic materials such as dictionaries, coursebooks (e.g., the Touchstone series; McCarthy, McCarten and Sandiford 2005) or vocabulary books (the English Vocabulary in Use series; McCarthy and O'Dell 2002). In the past, materials developers relied on their intuition when deciding on the content of coursebooks. However, as McCarthy (1998) notes, even native speakers are inaccurate at estimating the frequency of use of different linguistic elements. Therefore, at present, when large databases of language data have become available, lexicographers and materials developers search corpora and their findings inform what is included in dictionaries and coursebooks. It is worth mentioning that corpus analysis has led scholars to the idea of the lexical syllabus - an innovative approach to language teaching that directly uses corpus findings and organizes the content for teaching around frequent vocabulary. Sinclair and Renouf (1988) were the first authors who suggested the lexical syllabus following their work on the COBUILD project.



The International Corpus of the English Language

With the purpose to ensure the "comparability" of the text samples under consideration, the project for the creation of the International English language was primarily subordinated. Continuing the traditions of British descriptive linguistics, project manager S. Greenbaum has significantly expanded the traditional practice of registering English speech use, including new materials reflecting regional and dialectal characteristics. In conditions where countries that officially recognize English as a second state language are striving for "linguistic independence" and are considering their own version of the language as a criterion for correct use, it is important that the options do not diverge too far. The creation of an international database was intended to help preserve the identity of at least the written form of the English language.

REFERENCES

1. Coxhead, Averil 2000: A New Academic Word List. TESOL Quarterly 34(2).
2. Crystal, David 2003: English as a global language. Cambridge: CUP.
3. DeCock, Sylvie, Granger, Sylviane, Leech, Geoffrey and McEnery, Tony 1998
4. Gries, Stefan 2009, What is corpus linguistics? Language and Linguistics Compass.
5. Johns, Tim 1991, Should you be persuaded: two samples of data-driven learning materials. English Language Research Journal 4.
6. McCarthy, Michael 2001: Issues In Applied Linguistics. Cambridge: CUP.
7. McCarthy, Michael and O'Dell, Felicity 2004: English vocabulary in use. Cambridge: CUP.
8. McCarthy, Michael, McCarten, Jeanne and Sandiford, Helen 2005: Touchstone. Cambridge: CUP.