In linguistics, a **corpus** (plural corpora) is a large and structured set of texts (now usually electronically stored and processed). A corpus may be used to help linguists to analyze a language, or for the purpose of dictionary writing or language teaching. The British National Corpus (BNC) is a 100-million-word text corpus of samples of written and spoken English from a wide range of sources. The corpus covers British English of the late twentieth century from a wide variety of **genres** with the intention that it be a **representative** sample of spoken and written British English of that time.

- 1. What is a corpus?
- A. A type of large dictionary
- **B**. A single written text
- **C**. A tool for language analysis
- 2. Why was the BNC compiled?
- A. For the purpose of language teaching.
- B. To document written and spoken English from a particular period in time.
- **C**. To document the history of the English language.



Hoàn thành bảng từ vựng sau:	
Sort elements	
corpus	
genre	
representative	
đại diện, tiêu biểu	
tập văn, tập sao lục, văn thư	
thể loại	

