

## Spoken Corpus comes to life

- A** The compiling of dictionaries has been historically the provenance of studious professorial types bespectacled - who love to pore over weighty tomes and make pronouncements on the finer nuance of meaning. They were probably good at crosswords and definitely knew a lot of words, but the information they provided was always rather dry and dusty. The latest technology, and simple technology at that, is revolutionising the content of dictionaries and the way they are put together.
- B** For the first time, dictionary publishers are incorporating real, spoken English into their data. It gives lexicographers (people who write dictionaries) access to a more vibrant, up-to-date vernacular than any which has never really been studied before. In one project, 150 volunteers each agreed to discard their Walkman recorder to their waist and leave it running for anything up to two weeks. Every conversation they had was recorded. When the data was collected, the length of tapes was 35 times the depth of the Atlantic Ocean. Teams of audio typists transcribed the tapes to produce a computerised database of tens of thousands of words.
- C** This has been the basis - along with an existing written corpus - for the Language Activator dictionary described by lexicographer Professor Randolph Quirk as "the book the world has been waiting for." It tells advanced foreign learners of English how the language is really used. In the dictionary, key words such as "eat" are followed by related phrases such as "wolf down" or "be a picky eater", allowing the student to choose the appropriate phrase.
- D** "This kind of research would be impossible without computers," said Delia Summers, a director of dictionary publishers. "It has transformed the way lexicographers work. If you look at the word "like", you would intuitively think that the first and most frequent meaning is the verb, as in "I like swimming". It is in fact a preposition, as in: "she walked like a duck". Just because a word or phrase is used doesn't mean it should be in a dictionary. The sifting out process is as vital as ever. But the database does allow lexicographers to search for a word and find out how frequently it is used - something that could only be guessed at before.
- E** Researchers have found that written English works in a very different way to spoken English. The phrase "say what you like" literally means "feel free to say anything you want", but in reality it is used, especially by women, to show, by someone to prevent the other person voicing disagreement. The phrase "it's a question" is up on the database over and over again. It has nothing to do with enquiry, but it's one of the most common English phrases which has never been in a language learner's dictionary before: it is now.
- F** The Spoken Corpus computer shows how inventive and humorous people are when they are using language, by twisting familiar phrases for effect. It also reveals the power of the pauses and noises we use to save time, convey emotion, doubt and irony.
- G** For the moment, those benefiting most from the Spoken Corpus are foreign learners. "Computerised corpora allow lexicographers to search quickly through more examples of real English," said Professor Geoffrey Leech of Lancaster University. "They allow dictionaries to be more accurate and give a feel for how language is really used." The Spoken Corpus is part of the larger British National Corpus, an initiative carried out by a number of groups involved in the production of language learning materials: publishers, universities and the British Library.

## Questions 1-6

Reading Passage has seven paragraphs (A-G). Choose the most suitable heading for each paragraph from the list of headings below. Write the appropriate numbers (i-xi) in boxes 1-6 on your answer sheet. Paragraph C has been done for you as an example.

**NB There are more headings than paragraphs so you will not use all of them. You may use any heading more than once.**

### List of Headings

- i Grammar is corrected
- ii New method of research
- iii Technology learns from dictionaries
- iv Non-verbal content
- v The first study of spoken language
- vi Traditional lexicographical methods
- vii Written English tells the truth
- viii New phrases enter dictionary
- ix A cooperative research project
- x Accurate word frequency counts
- xi Alternative expressions provided

1..... Paragraph A

2..... Paragraph B

*Example*

Paragraph C Answer: xi

3..... Paragraph D

4..... Paragraph E

5..... Paragraph F

6..... Paragraph G

## Questions 7-11

The diagram below illustrates the information provided in paragraphs B-F of Reading Passage.

Complete the labels on the diagram with an appropriate word or words.

Use **NO MORE THAN THREE WORDS** for each space.

Write your answers in boxes **7-11** on your answer sheet

7.....

8.....

9.....

10.....

11.....

## Questions 12

Choose the appropriate letter **A-D** and write it in box **12** on your answer sheet

**12** Why was this article written?

- A To give an example of a current dictionary.
- B To announce a new approach to dictionary writing.
- C To show how dictionaries have progressed over the years.
- D To compare the content of different dictionaries

**Solution:**

- 1. vi    7. existing
- 2. ii    8. (related) phrases
- 3. x    9. meanings//forms
- 4. viii    10. spoken//real//oral
- 5. iv    11. noise//pauses//noises and  
pauses
- 6. ix    12. B