

---

UNIVERSITI SAINS MALAYSIA

Second Semester Examination  
Academic Session 2007/2008

April 2008

**HET 324 - COMPUTATIONAL LINGUISTICS**

Duration : 3 hours

---

Please check that this examination paper consists of TWO pages of printed material before you begin the examination.

Answer any FOUR questions.

1. Do you agree with the claim that language generation is more difficult than language analysis? Discuss.  
[25 marks]
  
2. According to Barnbrook (1996), the extent to which computational linguistics applications such as machine translation, lexicography, word processing, information retrieval, speech recognition and others make use of NLP varies greatly. In your opinion, which application would benefit users most if it has full NLP capability? Why? Discuss.  
[25 marks]
  
3. Between semantics and syntax, which linguistic aspect do you think is more important in language generation and analysis? Why?  
[25 marks]
  
4. Discuss how the corpus-based method of analysis facilitates the area of lexicography.  
[25 marks]
  
5. A computational lexicon may have as many as nine categories of information for one lexical entry. What are they? Using six of the categories, provide four different entries for the words *culture*, *plant*, and *date*.  
[25 marks]
  
6. What are the problems in machine translation that Barnbrook raises in his chapter? Use the following translations of Russian into English to explain the problems. Support your answer with further examples.

In the case of fire, do your utmost to alarm the hotel porter.  
Ladies are requested not to have children in the bar.  
Physician specializing in women and other diseases.  
Please leave your values at the desk (Hotel sign)

[25 marks]