

START OF QUIZ

Student ID:

**37105848, Yuen, Shing Hei
Robin**

Question 1

Topic: Lecture 3

Source: Lecture 3

Imagine that we have a parallel corpus (ie, a corpus containing sentences in two languages), and we want to extract a bilingual lexicon. What are some simple steps we could do to identify words that could be translations of each other? (2)

Question 2

Topic: Lecture 1

Source: Lecture 1

When would we *not* want to lowercase text prior to training a model? Give a concrete example. (1)

Question 3

Topic: Lecture 3

Source: Lecture 3

What properties of dictionaries make them an efficient choice for nesting complex lexicons.

(1)

Question 4

Topic: Lecture 2

Source: Lecture 2

Would a language with with lots of inflection have a higher or lower TTR than one with little inflection? Briefly explain. (1)

Question 5

Topic: Lecture 4

Source: Lecture 4

In class, we removed stopwords by using a lexicon. Can you think of another way that we could remove all closed class words? (1)

Question 6

Topic: Lecture 2

Source: Lecture 2

Do you think it's possible for a language not to follow a Zipfian curve? What consequences might that have on communication (if, let's say, if the curve were linear)? (2)

Question 7

Topic: Lecture 4

Source: Lecture 4

In French, negation is often indicated by "ne ... pas" (ie, "je ne parle pas" - "I am not speaking"; "tu ne conduis pas" - "You are not driving", etc.). However, in speech, one of the two is often dropped: "je ne parle." or "tu conduis pas.". Using this information, how would you determine whether a corpus was composed of written or spoken French? You don't need to write the code, but explain the logic that you would use to come to this conclusion. (2)

Question 8

Topic: Lecture 1

Source: Lecture 1

How would you reverse a string and keep the result? (1)

Question 9

Topic: Coding

Source: Coding

Write a function that determines whether the last consonant of the root of a verb has been doubled. For example: win -> winning. The function can take any string. What kind of error testing should you perform? (Hint - consonant doubling only occurs in certain environments). (3)

END OF QUIZ