1a) "To be or not to be, that is the question"

1-grams	2-grams	3-grams	4-grams
To	To be	To be or	To be or not
be	be or	be or not	be or not to
or	or not	or not to	or not to be
not	not to	not to be	not to be that
to	to be	to be that	to be that is
be	be that	be that is	be that is the
that	that is	that is the	that is the question
is	is the	is the question	
the	the question		
question			

1b)

Distinct 1-grams: 8
Distinct 2-grams: 8
Distinct 3-grams: 8
Distinct 4-grams: 7

1.c)
For a 15 word sentence the maximum number of n-grams is...
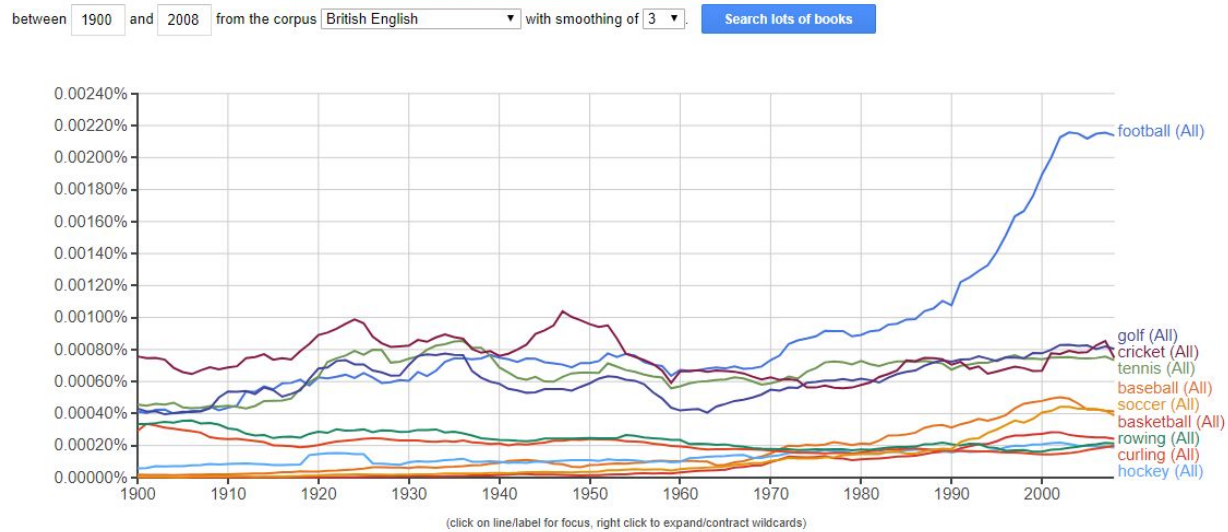2-grams: 14
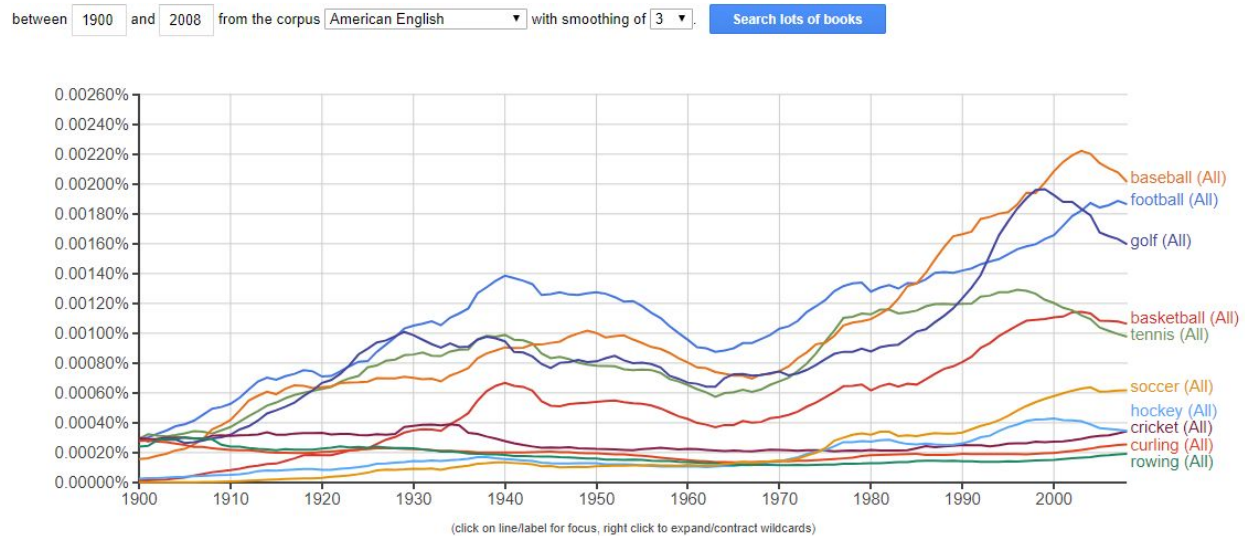5-grams: 11
7-grams: 9
17 grams: 0

2.a) In printed text from 1900 to 2008, "ketchup" appears more than "catsup" from the years 1977 to 2008.

2b) i.
British English corpus

American English corpus

ii.
Most popular sports in British English corpus 1950: Cricket, Football, Tennis
Most popular sports in American English corpus 1950: Football, Baseball, Golf

iii.
Most popular sports in British English corpus 2000: Football, Golf, Cricket
Most popular sports in American English corpus 2000: Baseball, Golf, Football
In the British English corpus, Football saw a meteoric rise in relevance in printed text rising far above golf by 2000. In the American English corpus, the most popular sports in 1950 had traded places by the year 2000.