Speech Signal Processing

Exercise 1

1. A corpus consists of the vocabular **A**, **B** and **C**. In the training the word **A** occurs 4 times, the **B** 3 times and the word **C** 0 times.

   (a) What is the occurrence probability of the sequence *ABABBCAAA-BAAB* if a unigram is used?

   (b) Compute the perplexity. Is it well defined?

   (c) Compute the unigram-probabilities using Jeffrey's method?

2. Language models: Good-Turing estimate
   The *Austen* test corpus has a vacabular size of $V = 14585$ words. It contains $N = 617091$ words and $\eta = 199252$ bigrams.
   In the table η_r denotes: number of bigrams which occured r-times in the corpus.

   (a) How many "unseen" bigrams do exist?

   (b) Compute and interpret r^* for $r = 0, 2, 7, 9, 843$.

r	η_r	r	η_r
1	138741	8	1342
2	25413	9	1106
3	10531	10	896
4	5997	...	
5	3565	843	1
6	2486	844	0
7	1754	...	

   (c) What is the bigram-probability of *she was* and *both sisters* without language model smoothing?

   (d) Compute the bigram-probabilities for the sentence *she was inferior to both sisters*. Use the table below and the Good-Turing estimate. (Assume that the predecessor of the word *she* is *person*).

w	$\#(w)$	$w_1 w_2$	$\#(w_1 w_2)$
person	223	person she	2
she	6917	she was	843
was	9409	was inferior	0
inferior	33	inferior to	7
to	20042	to both	9
both	317	both sisters	2

3. Speech production

(a) Describe the process of tone-production in humans.

(b) Draft a block model (Source-Filter) of speech production.

(c) The speech plots below show each a different sound. Where is the difference and what sounds (voiced, unvoiced) are produced?