1) a) Variations of words must be recorded.
By this, the number of words in the dataset increases a lot.
Variations are e.g. conjugates like
- eine
- einer
- eines
- ...

b) The prosody of a spoken sentence differs from single words.
Examples: - in a question, the last word is usually spoken in a higher pitch.
- The short breaks between words recorded each on its own makes the voice sounding machine like without any emotions.

2) $p(n) = \frac{1}{n \ln(1.78 \cdot N)}$

„der" $p(1) = \frac{1}{\ln(1.78 \cdot 10^5)} = 6,95\,\%$

„die" $p(2) = \frac{1}{2 \ln(1.78 \cdot 10^5)} = 3,47\,\%$

„und" $p(3) = \frac{1}{3 \ln(1.78 \cdot 10^5)} = 2,32\,\%$

„in" $p(4) = \frac{1}{4 \ln(1.78 \cdot 10^5)} = 1,74\,\%$

„den" $p(5) = \frac{1}{5 \cdot \ln(1.78 \cdot 10^5)} = 1,39\,\%$

$$\sum_{n=1}^{n_i} \frac{1}{n \ln(1.78 \cdot 10^5)} = 0,5$$

$$\sum_{n=1}^{n_i} \frac{1}{n} = \ln(1.78 \cdot 10^5) \cdot 0,5 = 7,1961$$

solution in python: 238 words

$$n_i = 238$$