Response to Data Frequency

```
p(d|\Lambda) = [0.5*p("text"|\theta_d) + 0.5*0.1]
           text the
                                                     x [0.5*p("the" | \theta_d) + 0.5*0.9]
                                       \rightarrow p("text" | \theta_d)=0.9 >> p("the" | \theta_d) =0.1!
                                         p(d'|\Lambda) = [0.5*p("text"|\theta_d) + 0.5*0.1]
           text the
                                                    x [0.5*p("the" | \theta_d) + 0.5*0.9]
d' = the the the ...the
                                                    x [0.5*p("the" | \theta_d) + 0.5*0.9]
                                                    x [0.5*p("the" | \theta_d) + 0.5*0.9]
    What if we increase p(\theta_B)?
                                                    x [0.5*p("the" | \theta_d) + 0.5*0.9]
```

What's the optimal solution now? $p("the" | \theta_d) > 0.1$? or $p("the" | \theta_d) < 0.1$?

Behavior 2: high frequency words get higher $p(w|\theta_d)$