

- Katz adjusted counts:

$$c_{katz}(w_{i-1}^i) = \begin{cases} d_r r & \text{if } r > 0 \\ \alpha(w_{i-1}) p_{ML}(w_i) & \text{if } r = 0 \end{cases}$$

- $\alpha(w_{i-1})$ is chosen so that $\sum_{w_i} c_{katz}(w_{i-1}^i) = \sum_{w_i} c(w_{i-1}^i)$:

$$\alpha(w_{i-1}) = \frac{1 - \sum_{w_i: c(w_{i-1}^i) > 0} p_{katz}(w_i | w_{i-1})}{1 - \sum_{w_i: c(w_{i-1}^i) > 0} p_{ML}(w_i)}$$

- Compute $p_{katz}(w_i | w_{i-1})$ from corrected count by normalizing:

$$p_{katz}(w_i | w_{i-1}) = \frac{c_{katz}(w_{i-1}^i)}{\sum_{w_i} c_{katz}(w_{i-1}^i)}$$