



Intuition: When we see a new *type* w in training, $++\text{count}(w)$; $++\text{count}(\text{novel})$
 So $p(\text{novel})$ is estimated as $T/(N+T)$, divided among $n_0 = Z$ specific novel types