

## Nikolas Zeitler

### 1.)

Pos	DE	JA	EN
PROPN	0.10837843	0.028879315	0.0663148
VERB	0.083962016	0.09164484	0.13372837
DET	0.124184385	0.0039868327	0.08289448
ADJ	0.0713481	0.018667493	0.06261822
NOUN	0.17868376	0.31341287	0.17184006
PUNCT	0.13095406	0.10860476	0.11720362
PRON	0.05055291	0.0039793598	0.083302595
PART	0.007383266	0.010398646	0.026778637
ADP	0.1085092	0.21149269	0.08508025
NUM	0.025589958	0.039924372	0.01927952
ADV	0.049996313	0.010230504	0.05117922
AUX	0.023179164	0.10919512	0.038252953
CONJ	0.02996895	0.010323917	0.032276418
SCONJ	0.006068897	0.03490627	0.0181729
X	0.0012406033	0.0043530082	0.011077974
$\Sigma$	1	1	1

### 2.)

The Entropies of the languages (computed with the formula in the Script) are:

**Germany** 3.4243083

**Japan** 2.922599

**England** 3.5823135

### 3.)

The KL Divergence of the languages (computed with the formula in the Script) are:

**Germany**  $\Rightarrow$  **Japan** 0.6898806

**Japan**  $\Rightarrow$  **Germany** 1.4371134

**Germany**  $\Rightarrow$  **England** 0.0799654

**England**  $\Rightarrow$  **Germany** 0.17958218

**Japan**  $\Rightarrow$  **England** 1.2869648

**England**  $\Rightarrow$  **Japan** 0.52966446

#### 4.)

##### German

Bigram	$P(t_1, t_2)$	$P(t_2 t_1)$	$PMI(t_2, t_1)$
DET NOUN	0.8611426	0.98808837	0.045320734
ADP DET	0.7680257	0.9198945	0.07792825
VERB PUNCT	0.9584749	0.996533	0.0033645257

##### English

Bigram	$P(t_1, t_2)$	$P(t_2 t_1)$	$PMI(t_2, t_1)$
DET NOUN	0.57754785	0.97919214	0.2807402
ADP DET	0.47009987	0.79531807	0.43125528
VERB PUNCT	0.7570088	0.9694145	0.13295418

##### Japanes

Bigram	$P(t_1, t_2)$	$P(t_2 t_1)$	$PMI(t_2, t_1)$
DET NOUN	0.10345173	1.0	0.013049663
ADP DET	0.10285143	0.104567185	0.015472563
VERB PUNCT	0.9209605	0.9998914	0.0027329659