

# Supplementary Material

**Manuel R. Ciosici**  
UNSILO A/S and  
Aarhus University  
Aarhus, Denmark  
manuel@cs.au.dk

**Leon Derczynski**  
IT University of  
Copenhagen  
Copenhagen, Denmark  
leod@itu.dk

**Ira Assent**  
Department of Computer Science  
Aarhus University  
Aarhus, Denmark  
ira@cs.au.dk

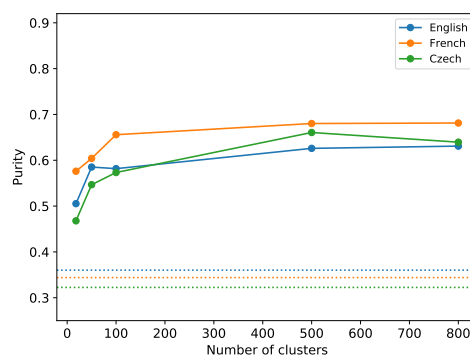
## 1 Experiments using Brown Clustering instead of Exchange

Brown clustering leads to lower AMI than Exchange. In this section we present *purity* and *AdjMI* scores on clusters derived using Brown instead of Exchange. We maintain the same setup as in the main paper, with the exception that we use Brown Clustering instead of Exchange. Since Brown Clustering outputs a hierarchical structure, we use only the flat clustering at the base of the hierarchy. This is also known as the point of peak Average Mutual Information.

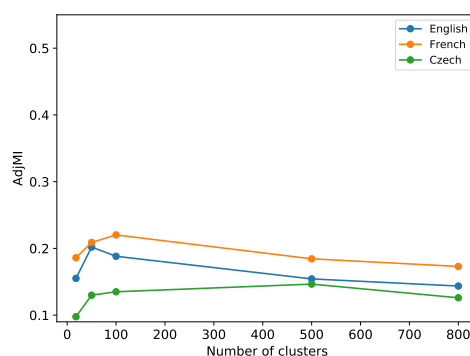
We set the active set parameter (the window) to  $k + 1$  for all experiments. Figures 1 and 2 in this document correspond to Figures 1 and 2 in the paper.

## 2 Correlation between AMI and morpho-syntactic information

For completeness, we include the correlation experiments for all combinations of  $k$  below. Tables 1 to 8 contain correlation results between AMI and *purity* and AMI and *AdjMI*. They follow the same setup as Tables 3 and 4 in the paper.

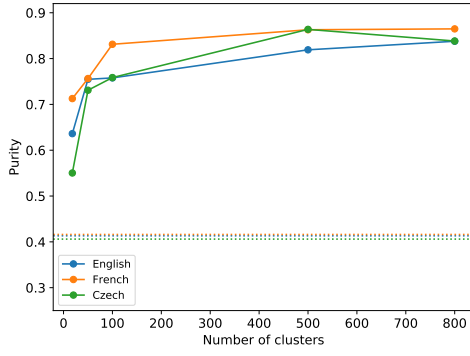


(a) Cluster purity. Dotted lines are baselines for  $k = 1$

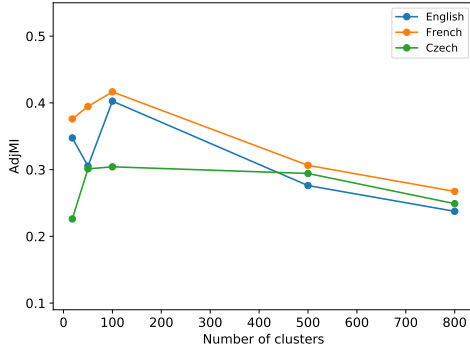


(b) Adjusted Mutual Information.

Figure 1: Cluster agreement with manual labels from UD. No frequency thresholding.



(a) Cluster purity. Dotted lines are baselines for  $k = 1$



(b) Adjusted Mutual Information.

Figure 2: Cluster agreement with manual labels from UD. Only words with frequency minimum 5.

Data Set	Pearson	Spearman
EN UD	0.9857	0.7173
FR UD	0.9822	0.8117
CZ UD	0.9884	0.7938

Table 1: Correlation between Average Mutual Information and PoS purity of the clustering resulted from Exchange with  $k = 50$ . Words with frequency  $< 5$  have been filtered.  $p < 0.01$  for all coefficients.

Data Set	Pearson	Spearman
EN UD	0.9893	0.5734
FR UD	0.9839	0.8415
CZ UD	0.9768	0.7938

Table 2: Correlation between Average Mutual Information and AdjMI of the clustering resulted from Exchange with  $k = 50$ . Words with frequency  $< 5$  have been filtered.  $p < 0.01$  for all coefficients.

Data Set	Pearson	Spearman
EN UD	0.9845	0.8758
FR UD	0.9877	0.7048
CZ UD	0.9797	0.8036

Table 3: Correlation between Average Mutual Information and PoS purity of the clustering resulted from Exchange with  $k = 100$ . Words with frequency  $< 5$  have been filtered.  $p < 0.01$  for all coefficients.

Data Set	Pearson	Spearman
EN UD	0.9843	0.8678
FR UD	0.9896	0.7069
CZ UD	0.9719	0.8102

Table 4: Correlation between Average Mutual Information and AdjMI of the clustering resulted from Exchange with  $k = 100$ . Words with frequency  $< 5$  have been filtered.  $p < 0.01$  for all coefficients.

Data Set	Pearson	Spearman
EN UD	0.9687	0.8893
FR UD	0.9849	0.8872
CZ UD	0.9800	0.9713

Table 5: Correlation between Average Mutual Information and PoS purity of the clustering resulted from Exchange with  $k = 500$ . Words with frequency  $< 5$  have been filtered.  $p < 0.01$  for all coefficients.

Data Set	Pearson	Spearman
EN UD	0.9607	0.8972
FR UD	0.9814	0.7285
CZ UD	0.9775	0.9742

Table 6: Correlation between Average Mutual Information and AdjMI of the clustering resulted from Exchange with  $k = 500$ . Words with frequency  $< 5$  have been filtered.  $p < 0.01$  for all coefficients.

Data Set	Pearson	Spearman
EN UD	0.9668	0.9176
FR UD	0.9803	0.8852
CZ UD	0.9715	0.9397

Table 7: Correlation between Average Mutual Information and PoS purity of the clustering resulted from Exchange with  $k = 800$ . Words with frequency  $< 5$  have been filtered.  $p < 0.01$  for all coefficients.

Data Set	Pearson	Spearman
EN UD	0.9518	0.8850
FR UD	0.9723	0.8848
CZ UD	0.9682	0.9416

Table 8: Correlation between Average Mutual Information and AdjMI of the clustering resulted from Exchange with  $k = 800$ . Words with frequency  $< 5$  have been filtered.  $p < 0.01$  for all coefficients.