

Data Generation Methods to Improve Recognizing Appropriate Taxonomic Word Relations

RUG-1-Pegasusers at SemEval-2022 Task 3



university of
 groningen

Frank van den Berg, Gijs Danoe, Esther Ploeger, Wessel Poelman, Lukas Edman, Tommaso Caselli

{f.l.van.den.berg, g.danoe, e.ploeger.1, w.g.poelman}@student.rug.nl

{j.l.edman, t.caselli}@rug.nl

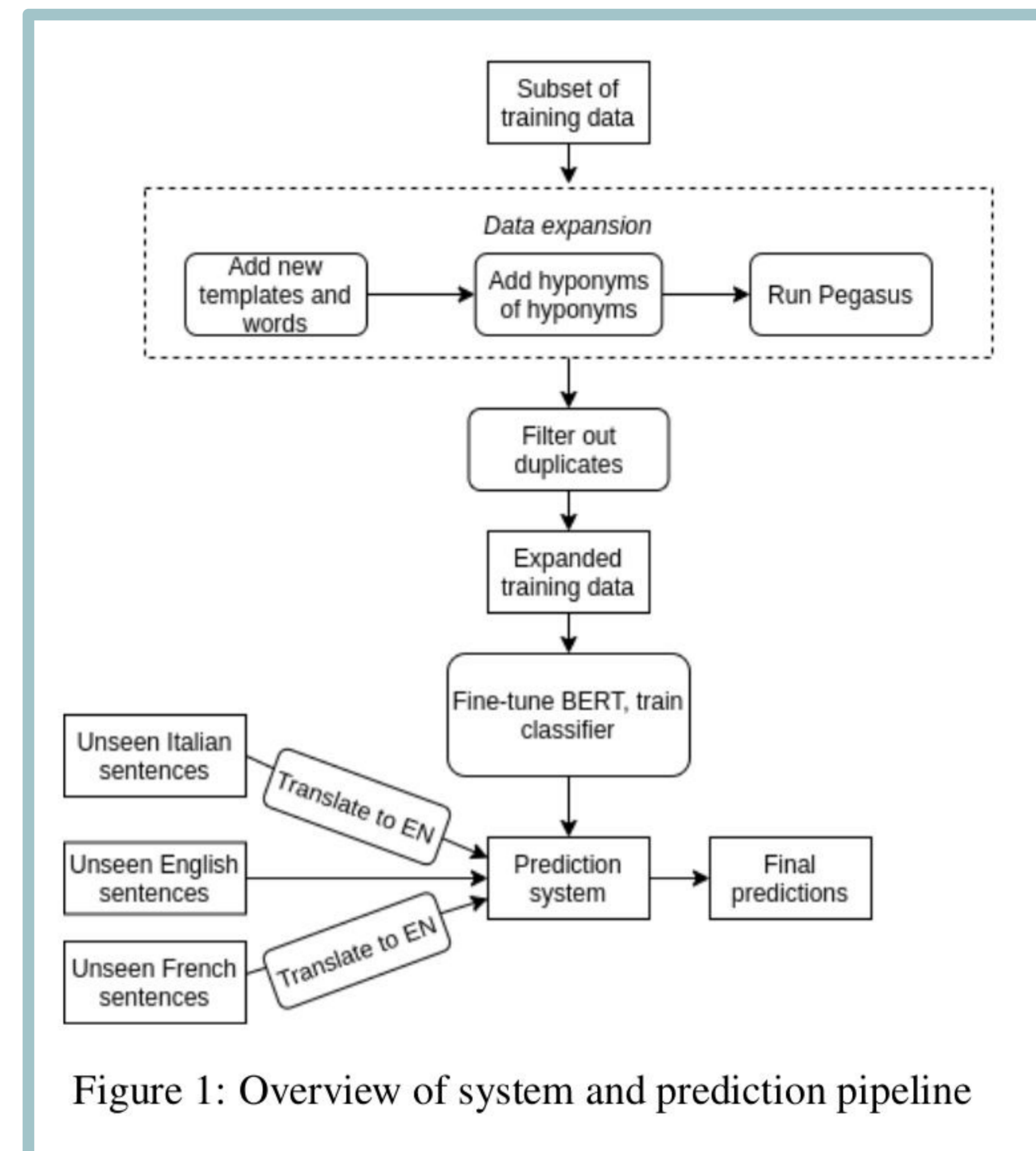
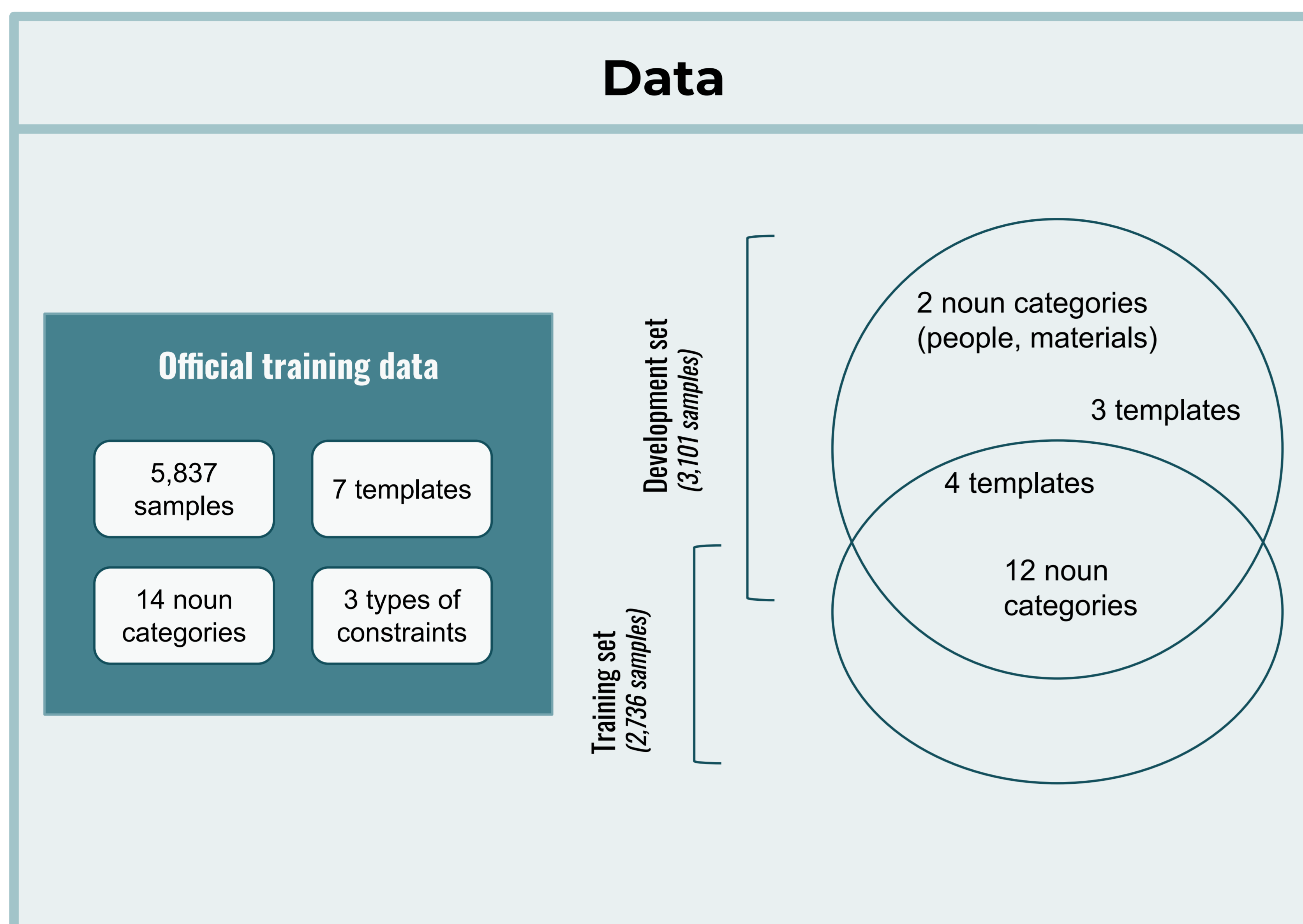
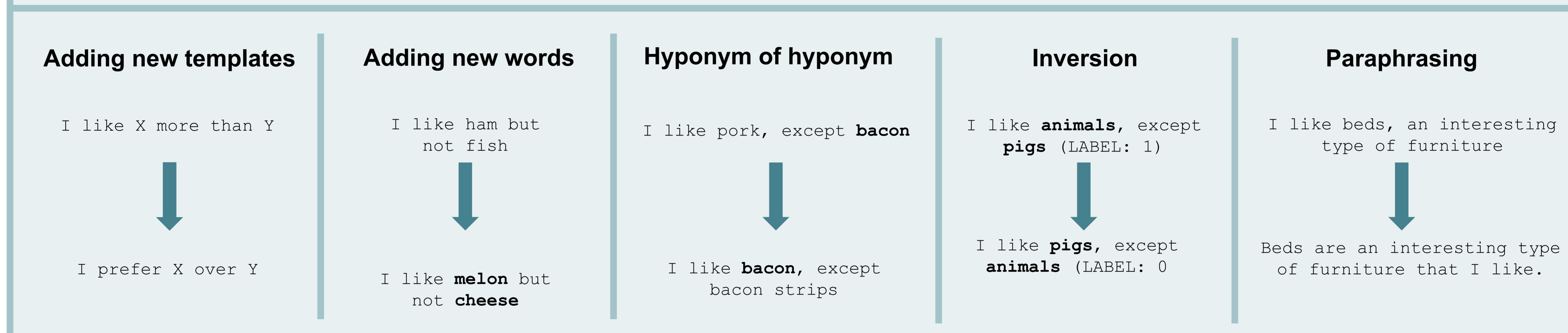


Figure 1: Overview of system and prediction pipeline

Data generation techniques

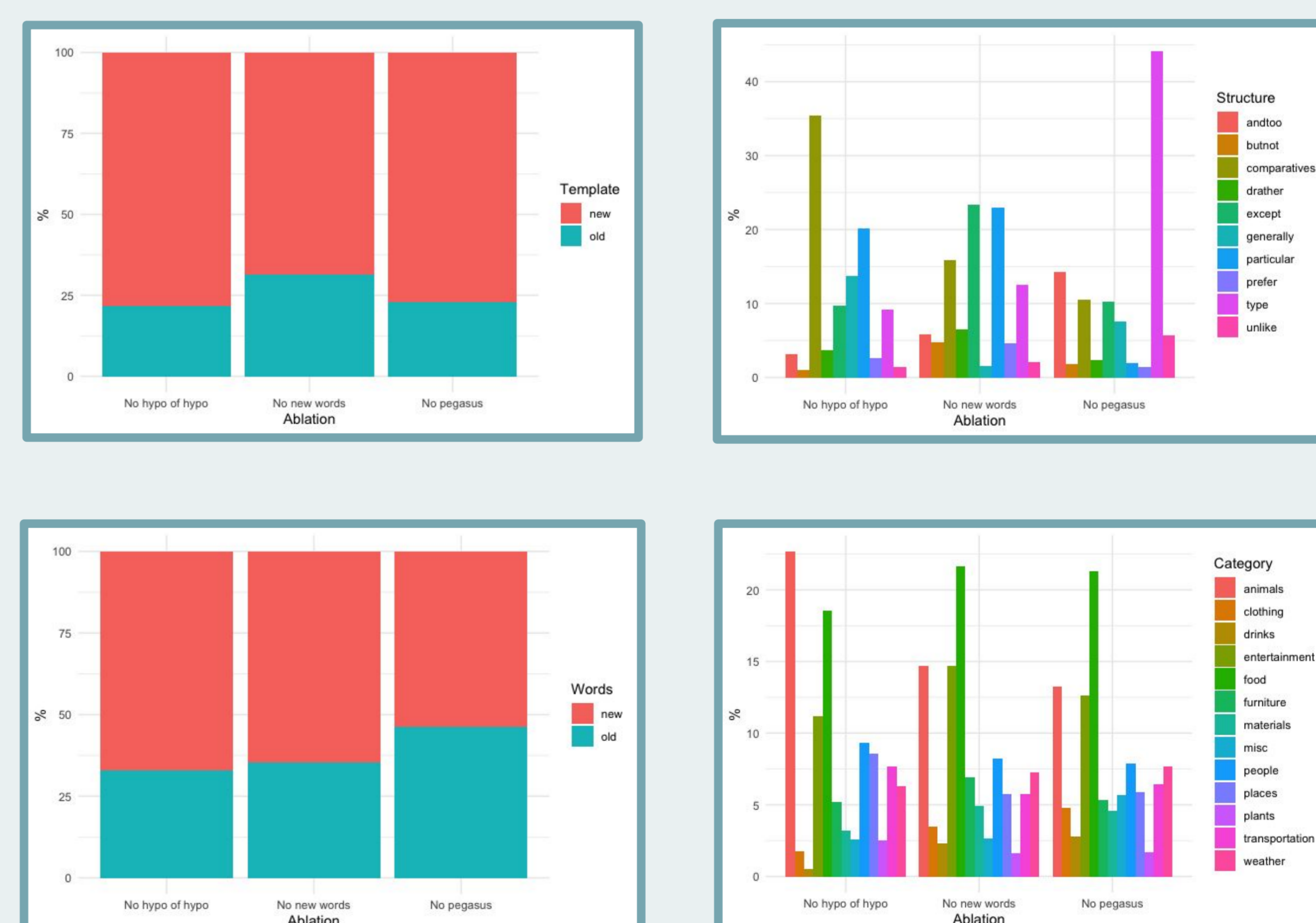


Results

Language	Precision	Recall	Macro F1
Italian	83.42	69.25	75.682
French	84.21	70.96	77.019
English	86.07	70.44	77.474

Results on official test set

Error Analysis



Conclusions

- All methods enable better performance
- Especially adding new templates, hyponyms of hyponyms, new words and paraphrasing are effective
- Using inversion is not that effective

Future work:
What causes these differences?