Testing the role of metadata in metaphor identification
Egon Stemle (EURAC Research, Bolzano), Alexander Onysko (University of Klagenfurt)

Background: Previous research has shown a positive information gain when using word embeddings from learner corpus data for metaphor classification in a neural network (Stemle & Onysko, 2018)
Aim: Explore the potential influence of the data structure in the annotated part of the ETS Corpus of Non-Native written English; particular focus on: proficiency ratings, essay prompt and L1 of the learner
System: fastText word embeddings from different corpora in a bi-directional recursive neural network with long-term short-term memory (LSTM BiRNN); a flat sequence to sequence neural network with one hidden layer using TensorFlow+Keras (Abadi et al., 2015) in Python.

**Table 1** illustrates a numerical bias of metaphors per prompt. Three groups emerge: P1-P3 highest rate; P4-P6 lowest rate; P7 and P8 medium rate → significant difference btw the groups of highest and lowest scoring prompts.

**Table 2** explains the reasons for the higher rate of metaphors in P1, P2, and P3. These prompts contain and trigger certain metaphors that are picked up by the learners in their prompted argumentative essays while this is not the case in the other prompts.

**Conclusion:** Table 3 indicates that the features of the metadata that showed a significant bias for the number of metaphors in the ETS data (prompt and proficiency) do not influence the performance of the system. Among the other metadata, text ID even shows a negative effect on the identification rate.

| # words | met. typ/tok | % met. per w | mean # met. |
|---------|--------------|--------------|-------------|
| P1 | 8059 | 0.57 | 4.5 | 15.70 |
| P2 | 7493 | 0.60 | 4.4 | 15.71 |
| P3 | 7947 | 0.59 | 4.8 | 17.23 |
| P4 | 8076 | 0.84 | 2.1 | 7.52 |
| P5 | 8455 | 0.83 | 2.4 | 8.96 |
| P6 | 8446 | 0.71 | 2.2 | 7.83 |
| P7 | 7516 | 0.70 | 3.2 | 11.05 |
| P8 | 7923 | 0.76 | 3.3 | 11.82 |

**Table 1:** Number of metaphors per prompt in the annotated ETS training set.

**Table 2:** Most frequent metaphorical expressions in P1, P2, and P3.

| metaphorical expression | # of occ. | metaphorical expression | # of occ. | metaphorical expression | # of occ. |
|-------------------------|-----------|-------------------------|-----------|-------------------------|-----------|
| P1 "broad(er) knowledge" | 60 | have/-ing/has "~ time" | 55 | give/-s/-ing/-en "~ time/help" | 29 |
| P2 focus/-ed "~ on a particular subject/field" | 12 | spend/-s/-ing "~ time/ hours/years" | 14 | have/-ing/has "~ time" | 24 |
| P3 * give/-s/-ing | 10 | * face "~ problems/change/responsibilities" | 6 | spend/-ing/-t "~ time" | 16 |

**Table 3:** 10-fold CV comparison of training with different metadata on the TOEFL dataset.

| | Verbs | All-POS |
|---|-------|---------|
| Baseline | Pr 0.53 (+/- 0.04) | All-POS 0.53 (+/- 0.04) |
| | Re 0.64 (+/- 0.07) | 0.63 (+/- 0.05) |
| Prompt | F1 0.57 (+/- 0.02) | 0.57 (+/- 0.03) |
| | Pr 0.50 (+/- 0.06) | 0.51 (+/- 0.05) |
| | Re 0.66 (+/- 0.10) | 0.64 (+/- 0.05) |
| | F1 0.56 (+/- 0.02) | 0.56 (+/- 0.03) |
| Proficiency | Pr 0.51 (+/- 0.07) | 0.51 (+/- 0.05) |
| | Re 0.68 (+/- 0.09) | 0.65 (+/- 0.07) |
| | F1 0.57 (+/- 0.02) | 0.57 (+/- 0.04) |
| Text ID | Pr 0.42 (+/- 0.08) | 0.48 (+/- 0.08) |
| | Re 0.72 (+/- 0.08) | 0.60 (+/- 0.10) |
| | F1 0.52 (+/- 0.05) | 0.52 (+/- 0.03) |