Merkel had already brought up a limited solidarity fund for discussion.

Merkel hatte bereits einen befristeten Solidaritätsfonds ins Gespräch gebracht.

1. Dependency parsing with the MATE parser [2]

2. Extraction of subject-verb dependencies using the SubCat-Extractor [7]

3. Selection for annotation: verb-dependent pairs with nouns as dependents

Merkel hatte bereits einen befristeten Solidaritätsfonds ins Gespräch gebracht.

- Annotation guidelines (condensed, original in German):
  1. Is the verb-dependent pair correct?
     - Yes → proceed to compositionality scoring in step 2
     - No → annotate “ERROR” and comment
  2. What compositionality score do you assign to the pair?
     - Composition: 0 1 2 3 4 5 (non-compositional)
  3. Supporting tests (adopted from the PARSEME shared task annotation guidelines):
     - CRAM: The verb-dependent pair contains a cranberry word.
     - LE4: Replacing one component with a semantically similar expression yields unexpected grammatical/semantic shifts.
     - VF4: The verb contributes almost nothing to the overall expression.
     - Yes → choose a score between 1 and 5
     - No → choose 0 as a score

- Annotators: Rafael Ehren & Lea Vanessa Möllmann (Düsseldorf), Alexander Frey, Daniel Helbig, Gjirianna Jagfeld & Florian Lux (Stuttgart)

- Evaluation based on 3 full annotations
- 25,317 judgments per annotator
- 2,547 verb-dependent pairs were annotated with ERROR by at least one annotator

- Inter-annotator agreement (IAA):
  - Mean Spearman’s $\kappa$: 0.416
  - Fleiss’ $\kappa$: 0.354 (binarized with [0.2, 1.4, 3])

- Cohen’s Kappa ($\kappa$):
  - CLEAN-BNC: 0.322
  - CLEAN-MORPH: 0.470

- Krippendorff’s Alpha ($\alpha$):
  - CLEAN-BNC: 0.412
  - CLEAN-MORPH: 0.68

- Distribution of scores:
  - Highly skewed toward 0 (compositionality)
  - Very strong baseline for computational models

- Further annotations (ongoing)
- Cleansing (planned): e.g. majority vote
- Shared task at KONVENS 2018 (planned): prediction of degree of compositionality

References