

Annotation Guidelines

for

German Emotion Stimulus Detection

We use this guidelines for annotating stimulus events in German news headlines. The annotation procedure is split in two phases: **Phase 1** annotation of emotion classes and the existence of emotion cue words or experiencers. **Phase 2** token-level annotation for stimulus events. This guidelines serve to help annotators by their decisions and provide short explanations for *experiencer*, *cue* and *stimulus*, used emotion set as well as the concept of sequence labeling using IOB scheme. For this task, various questions are formulated, which need to be addressed during annotation phases.

1 Guidelines

Terms:

1. **Experiencer:** An entity or a person who experiences an emotion.
2. **Cue:** A word that might have an emotional meaning or indicates an emotional content (e.g. “wütend”, “Unfall”, “fröhlich”).
3. **Stimulus:** An event, a person or a circumstance that evokes an emotional response in the experiencer.

EMOSET

In this EMOSET we will define ten different emotion classes. The numeric indication of these emotions in parentheses will be used for annotation.

1. Happiness (1) should be annotated, when the text describes a gratifying event. Cue terms for this emotion could be “glücklich”, “erfreut”, “lachen” or “Freude”.
2. Sadness (2), if a person is experiencing a sad event within the text. Also, if the text contains keywords such as: “tot”, “Trauer”, “deprimiert” or “unglücklich”.
3. Fear (3) is annotated, if the text mentions concerns or fears of the experiencer or has terms like “ängstlich”, “fürchterlich” or “zweifelnd”.
4. Disgust (4) is annotated, when the text describes an unpleasant event or the experiencer is disgusted. Also, keywords such as “widerlich”, “Abneigung”, “Unmensch” or “Schande” could indicate this emotion.
5. Anger (5) is to be annotated if the text hints an event that annoys the experiencer. Possible keywords for this emotion are “Kritik”, “verärgern”, “sauer” or “Gewalt”.
6. Positive Surprise (6) should be annotated if the text mentions a surprising event, that does not have negative effect on experiencer.

7. Negative Surprise (7) is to be annotated if a surprise is mentioned that could have negative impact or outcome for experiencer.
8. Shame (8) is to be annotated, when the text describe an event, which triggers shame in experiencers. Further, the text has keywords such as “beschämend”, “Schamhaftigkeit” or “bereuen”.
9. Hope (9) is should be annotated, when an event stimulates a hopeful feelings in the experiencer. Also, when the text include keywords such as “hoffen”, “erwarten” or “versprechen”.
10. Other (10) should be annotated, when the expresses an emotion that is not included in this set.

IOB scheme:

The task of detecting German emotion stimuli is defined as sequence labeling. Therefore, we use the IOB scheme for annotating token span that describes the cause of an emotion in the text.

- B: Begin of the stimulus span
- I: Inside of the span
- O: Outside the stimulus span

Linguistics Rules for Annotating Emotion Stimulus

1. Stimuli can be described by verbal or nominal phrases.
2. Subjunctions like “because of” belong to the sequence.
3. Conjunctions like “and”, “or” and “but” connect main clauses. They can therefore belong to a stimulus sequence.
4. Antecedents, if present, are annotated as stimuli.
5. If antecedent is not present, an anaphora may be annotated instead.
6. Composites with “-” are considered as a single word.
7. Stimuli can include several tokens or only one word.
8. Punctuations should not be labeled as stimulus. However, they might belong to the cause phrase as long as the stimulus is mentioned by multiple main clauses.

Further Rules

1. Stimuli could be described by one or multiple tokens
2. Colons or semicolon might be usually preceded by the subject of the headline. Stimuli are therefore often found after those special characters. However, that does not exclude that the subject of the headline can also cause an emotion.
3. Quotes could also describe the cause event of an expressed emotion. In such cases, we do not include quotation marks in stimulus spans. Further, we assume that only a certain part of the quote describe an expressed emotion.

2 Questionnaires

2.1 Phase 1: Emotion Annotation

Question 1: Are there terms in the headline which could indicate an emotion?

If:

- Yes: annotate with 1
- No: annotate with 0

Question 2: Does the text specify a person or an entity experiencing an emotion?

If:

- Yes: annotate with 1
- No: annotate with 0

Question 3: Which emotion is most provoked within the text?

- Use emotion from EMOSSET: Annotate the emotion with respect to cue or (and) experiencer annotations.
- Only consider the most dominant emotion expressed by the text. Therefore, do not annotate multiple emotion categories.
- Note that *no emotion* or 0 should be annotated if no cue and no experiencer exist in the sentence.

2.2 Phase 2: Stimulus Annotation

Question 4: Which sequence of tokens describe an event, a person or a circumstance that triggers an emotion?

- Use IOB scheme (explained above) to annotate emotion stimulus segments

3 Examples

Following examples should help annotators to understand, what is the goal of this annotation and what needs to be annotated. Also, we include few headlines in this section that might not contain emotional contents.

Annotation Examples

1. Tarifrunde im öffentlichen Dienst gefährdet neue Grundrechte (FAZ)

- Q1: 1 (“gefährdet”)
- Q2: 0
- Q3: 3 (*fear*)
- Q4: B I I I O O O O (“Tarifrunde im öffentlichen Dienst”)

2. Millionen-Hilfsprogramm der Bundesregierung: Neue Hoffnung für Düsseldorfer Kommödchen und Co.(RP-Online)
 - Q1: 1 (“Hoffnung”)
 - Q2: 1 (“Düsseldorfer Kommödchen und Co.”)
 - Q3: 8 (*hope*)
 - Q4: B I I O O O O O O O (“Millionen-Hilfsprogramm der Bundesregierung”)
3. Friedenspreis des Deutschen Buchhandels für Inder Amartya Sen
 - Q1: 0
 - Q2: 1 (“Amartya Sen”)
 - Q3: 1 (*happiness*)
 - Q4: B I I I O O O O (“Friedenspreis des Deutschen Buchhandels”)
4. UN erhalten Hunderte Millionen Euro für Rohingya-Flüchtlinge
 - Q1: 0
 - Q2: 1 (“UN”)
 - Q3: 6 (*pos. surprise*)
 - Q4: O O B I I I I (“Hunderte Millionen Euro für Rohingya-Flüchtlinge”)

Headlines without emotional content

1. Top /Facts

- “Kennst du schon diese 10 Fakten über...”
- “Sag uns was du zum Frühstück isst und wir sagen Dir wie dein Leben in 5 Jahren aussieht”

2. General Reports

- “75. Geburtstag: So verbringt Ralph Siegel den Herbst seines Lebens”
- “Coronavirus-Pandemie: Im Bundestag gilt ab Dienstag Maskenpflicht”
- “Warum junge Konservative Trump wählen”

3. Tips

- “10 hilfreiche Methoden: Wasser im Ohr loswerden”
- “Das neue Huawei”