

Guidelines for the Task of Annotating Sentiment Views Conveyed by Multiword Expressions

The expressions to be annotated are multiword expressions (MWEs), e.g. *pull the plug*, *keep up with the times*, *add fuel to the fire*, *put out feelers* or *set in stone*. Each MWE consists of several word tokens (e.g. *pull the plug* consists of *pull*, *the* and *plug*). Often, the meaning of the MWE does not reflect the meaning of its component word tokens. For instance, *pull the plug* is meant metaphorically and has a meaning similar to the verbs *terminate* or *stop*. In this annotation task, annotators should exclusively consider the meaning of the MWE as a whole. The meaning of the individual word tokens should **not** be considered.

Each MWE conveys some subjective information. Therefore, these MWEs can be considered a subset of opinion expressions. Opinion expressions typically also convey a sentiment view. By sentiment views, we understand the perspective of the holder of some opinions. There are two different perspectives: On the one hand, the holder can be the implicit speaker of the utterance. For example, in

(1) *Minecraft is a game that always **keeps up with the times**.*

the implicit speaker of the utterance has some positive sentiment towards the computer game *Minecraft*. The positive sentiment is conveyed by the MWE *keep up with the times*. This form of sentiment view is referred to as **speaker view**. On the other hand, an opinion expression may also convey sentiment of the entities participating in an opinion event. For example, in

(2) *Peter **has a soft spot** for the French cuisine.*

the holder *Peter* has a positive sentiment towards *the French cuisine*. This positive sentiment is conveyed by the MWE *have a soft spot*. This form of sentiment view is referred to as **actor view**.

In this annotation task, we just focus on the binary distinction between actor view and speaker view. That is, for actor views, we do not specify the specific (syntactic) realization of the opinion holder (i.e. is the holder the agent of the MWE, is it the patient of the MWE, etc.).

Of course, speaker and actor view need not always convey positive sentiment. Instead, they may also convey negative sentiment:

(3) *His latest remarks only **added fuel to the fire**. (speaker view)*

(4) *Twitter **took action against** extremist tweets. (actor view)*

or neutral sentiment:

(5) *Events **come to a climax** on the night of the election results, on March 19. (speaker view)*

(6) *The Prime Minister is about to **make a statement** on the new health plans of the government. (actor view)*

This has no impact on the choice whether to tag an MWE as conveying an actor view or a speaker view.

Further hints:

1. Although this is an out-of-context annotation, i.e. MWEs are annotated in isolation rather than in the context of a specific sentence, annotators may consult an online dictionary, e.g. www.macmillandictionary.com, in order to learn about the typical use of some MWE. In particular, it may help to consider the example sentence(s) that are offered in addition to the definition of an MWE.
2. In deciding whether an MWE conveys an actor or speaker view, annotators should try to figure out whether it is one of the persons mentioned in a possible example sentence whose sentiment is conveyed by the given MWE. In addition, these persons must be **dependents** of the MWE, i.e. agent or patient. For example, in:

(7) *Peter thinks that Mary **has a great future** ahead of her.*

Peter has some positive sentiment towards *Mary* but that sentiment is not conveyed by the MWE *have a great future*. The noun phrase *Peter* is not a dependent of the MWE but the other opinion expression *think*. Therefore, the MWE actually conveys a speaker view.

3. There may be MWEs which have more than one meaning. Typically, this is reflected in a dictionary entry of the MWE having several senses. For this annotation task, we always assign exactly one label for each MWE. In other words, our annotation is on the lemma level rather than on the sense level. For MWEs with multiple senses, we have the following advice:

1. If all senses convey the same sentiment view, then this poses no problem. It is possible to assign exactly one sentiment view to the given MWEs.
2. If the senses convey different sentiment views, then the annotator needs to choose one of the given senses. We suggest the following prioritization:

- If there is one sense with a literal meaning and one with a figurative meaning, the sense with the figurative meaning should be chosen. Literal senses of MWEs typically do not convey subjectivity. However, we only consider MWEs with a subjective connotation in this work.
- Annotators should prefer a more general sense over a specialized sense. This is typically reflected in the order in which the senses of an MWE appear in a dictionary. The first entry is the more frequent sense and it should be preferred (provided it has a figurative meaning with subjective connotation, otherwise the second sense should be taken). Moreover, more specific senses are also often indicated by their domain use (e.g. *technical*, *medical*, *sports* etc.).

3. Example: dictionary entry for **take a back seat**:

- (sense 1) to deliberately become less active, and give up trying to control things: *I'll be happy to **take a back seat** when Robin takes over.*
- (sense 2) to become less important: *Other issues must **take a back seat** to this crisis.*

"Sense 1" conveys an actor view while "sense 2" conveys a speaker view. Since "sense 1" is the first sense and thus more general; it should be preferred for this lexical entry.

4. There may be cases in which the same MWE conveys both actor and speaker view despite the fact that it only has one sense. In these cases, annotators are asked to annotate the sentiment view that they think is more prominent for the given MWE. Example:

(8) *Peter tries his best.*

The MWE conveys a speaker view since the speaker of that utterance evaluates Peter's behaviour. It has a positive sentiment towards him since it recognizes Peter's genuine efforts. On the other hand, Peter also has some positive sentiment regarding an implicit patient, i.e. some goal he wants to achieve. So, this would also qualify for an actor view. For this MWE, we would recommend to label the MWE as conveying a speaker view since it is more prominent than the actor view. Typically, the goal of *tries one's best* is not realized as a dependent of the MWE (i.e. it remains an implicit patient).

This clearly indicates that the resulting actor view is not that prominent.

5. One may get the idea that if the given MWE has a figurative meaning, e.g. *reinvent the wheel*, *hit the ceiling* or *add fuel to the fire*, then the MWE conveys some speaker view. This may often be the case but it is not always true. In other words, there are figurative MWEs, such as *pat on the back* or *take one's hat off*, that actually convey an actor view. Therefore, annotators should not annotate MWEs that are used with their figurative reading with a bias on speaker views but consider each MWE separately.
6. One may also get the idea that if the given MWE is a light-verb construction, e.g. *have faith*, *take revenge* or *wage war*, then the MWE conveys an actor view. This may often be the case but it is not always true. In other words, there are light-verb constructions, e.g. *have guts* or *make trouble*, that actually convey a speaker view. Therefore, annotators should not annotate light-verb constructions with a bias on actor views but consider each MWE separately.