

Emotional Factuality Annotation Guidelines

by Samuel Haymann, Lizzie Liang, and Erik Andersen

[Introduction](#)

[Corpus](#)

[Annotation Tasks](#)

[Event](#)

[Events denoted by verbs](#)

[Events denoted by nouns](#)

[Events denoted by adjectives](#)

[Conjunctions and Lists](#)

[Embedded Clauses](#)

[Multi-word Expressions](#)

[Phrasal verbs](#)

[Idioms](#)

[Ellipsis](#)

[Source](#)

[Source mentioned in text](#)

[Author as source](#)

[Quotes](#)

[Retweets](#)

[More than one source](#)

[Confidence](#)

[Positive, High Confidence](#)

[Positive, Low Confidence](#)

[Uncertain/Neutral](#)

[Negative, Low Confidence](#)

[Negative, High Confidence](#)

[Emotion](#)

[Positive](#)

[Negative](#)

[Neutral](#)

[Links Between Annotations](#)

[Annotation Example](#)

Introduction

This document provides guidelines on annotating text with information related to the factuality of events and emotion present in the text.

Corpus

For this annotation exercise, we will be using Trump's tweets as a corpus. Trump's archived tweets are available as part of a corpus at <http://www.trumptwitterarchive.com/>. This site contains all tweets made by the twitter handle @realdonaldtrump, and is updated in real time.

Annotation Tasks

Event

The first task is to select the event or events in the text. We are defining an event as the main predicate in the sentence that can be tied to a situation that has happened, or a circumstance that holds true. Each sentence should have one main event predicate, unless there is a conjunction (e.g. *and*, *but*, *therefore*, *because*) in which case there can be more than one. Events in this annotation project will be a subset of the work done in TimeML¹ and will be closer to the 'root' found in Universal Dependencies². For the most part, this event will be represented by a verb, but in some cases it may also be represented by a noun or adjective. In the case of nouns and adjectives, they will be predicative complements or predicative adjectives, meaning they will follow predicates such as copulas (e.g. *be*), basic causative verbs (e.g. *make*, *do*) or perception verbs (*seem*, *look*). Copulas, basic causative verbs, and perception verbs should *not* be annotated as events.

There may also be texts with no events, such as "Happy New Year!" where no annotation should be added.

Some portion of the text may be cut off because of its length, in which case only a sentence fragment will be visible. In this scenario, if an event is not easily discernible, the fragment should not be annotated.

¹ <http://timeml.org/tempeval2/tempeval2-trial/guidelines/EventGuidelines-050409.pdf>

² <https://universaldependencies.org/guidelines.html>

In the examples below, all events to be annotated are highlighted in red. However, only the ones related to the sections they are in are **bolded** with red color (all other events are indicated with *italics*). Additionally, each event is followed by angle brackets to reflect the annotation that will be added.

Events denoted by verbs

For an event denoted by a verb, only the main verb will be annotated

1. [Trump] alt was announced today by the U.S. Treasury that additional large scale Sanctions would be **added**<event1> to those already existing Sanctions on North Korea. I have today **ordered**<event2> the withdrawal of those additional Sanctions!
2. [Trump] ISIS **uses**<event1> the internet better than almost anyone, but for all of those susceptible to ISIS propaganda, they are now being **beaten**<event2> badly at every level....

Events denoted by nouns

These will mostly be predicate nominatives or verbal nouns, as shown below.

3. [Trump] "Our own Benjamin Hall is doing fantastic **reporting**<event1> on ISIS right on the from line (True). ISIS was willing to *die*<event2> but now, because of big pressure, save for a few people in caves, most have *surrendered*<event3>. A testament to our President." Thank you Pete Hegseth @foxandfriends
4. [Trump] It is my pleasure to announce that @StephenMoore, a very respected Economist, will be nominated to *serve*<event1> on the Fed Board. I have known Steve for a long time – and have no doubt he will be an outstanding **choice**<event2>!

Events denoted by adjectives

These adjectives will follow a predicate, such as "seems" or "looks," in a number of cases, in the same sense as is used below.

5. [Trump]Secretary of Defense Chuck Hagel seems so **lost**<event1> and, frankly, **dumb**<event2>. He can't even speak properly. Poor leader in these very dangerous times!
6. [Trump]The White House never looked more **beautiful**<event1> than it did returning last night. Important meetings taking place today. Big tax cuts & reform.

Conjunctions and Lists

Annotate more than one event when two predicates are separated by a conjunction (like *but*, *and*, *because*, *therefore*, *although*), such as can be seen in the example below. Note that in some cases (notably with conjunctions like *while* and *although*), the conjunction

can appear before the first predicate, and therefore acts more as a “trigger word” than a separator. This can be seen in sentence 8 below, with the word *although*. Finally, annotate more than one event when two or more predicates are separated by commas, such as is seen when listing predicates (as can be seen in sentence 9 below).

7. [Trump] ISIS **uses**<event1> the internet better than almost anyone, **but** for all of those susceptible to ISIS propaganda, they are now being **beaten**<event2> badly at every level....
8. [Trump] @AlanDersht, a brilliant lawyer, who **although** a Liberal Democrat who probably didn't **vote**<event1> for me, has **discussed**<event2> the Witch Hunt with great clarity and in a very positive way. He has *written*<event3> a new and very important book.
9. [Trump] RT @WhiteHouse: All across America, hardworking taxpayers **balance**<event1>: their household budgets, **find**<event2> ways to do more with less, **and** still **manage**<event3>...

Embedded Clauses

The main event will typically be in an embedded clause headed by a verb when present. For example, in sentence (e) *serve* is the main event, not *nominated* or *announced*.

10. [Trump] It is my pleasure to announce **that** @StephenMoore, a very respected Economist, will be nominated **to serve**<event1> on the Fed Board. I have *known*<event2> Steve for a long time – and have no doubt he will be an outstanding *choice*<event3>!

Multi-word Expressions

Phrasal verbs

Only annotate the verbal part, not its particle:

11. [Trump] Perhaps a new meeting will be **set**<event1> **up** with the @nytimes. In the meantime they continue to *cover*<event2> me inaccurately and with a nasty tone!

Idioms

In case of idiomatic phrases, annotate the whole phrase as the event

12. [Trump] @EdwardFCox and @RobAstorino, who I like, are *intent*<event1> on another New York State suicide mission. They've got to **get their act together**<event2>!

Ellipsis

In case of VP ellipsis, only annotate the event that is represented by the event in the sentence

13. [Trump] RT @WhiteHouse: The American dream is back because President Trump **understands**<event1> what many experts don't: that restoring American greatness i...

Source

The second task is to identify the source of the event. Factuality can be seen as a matter of perspective, so recognizing the source (or sources) of a piece of information is crucial for this project. The source of the event is defined as, at a particular point of time, one or more participants who report on an event with a certain level of confidence (see the section [Confidence](#) below). Different sources can have very different confidence levels with respect to the same event.

Sources are cognitive individuals who hold a specific stance regarding the factuality status of events in text. They correspond to the following types:

- The author of the text (speaker or writer).
- Other sources mentioned in the text. These sources are explicitly introduced in discourse by words such as: *say, think, according to, know, see, announce, or pretend*. Similarly, impersonal constructions (it seems, it is clear, ...) or passive constructions with no agentive argument (it is expected) should be attributed to the author.

Each event should have one source at the same point of time. In the case of more than one source present in the same event, the source in the text should be selected if the source is not the author (Trump).

In the following section, events mentioned in discourse always have an implicit source, namely, the author of the tweet. For retweets the source should be the author of the original tweet.

Notation: In the examples below, events will continue to be marked in red text, and the source will be marked with blue text. Additionally, for clarity's sake, we have noted the annotation that each carries using angle brackets and used numbers to represent relationships. The links representing relationships will be discussed in greater detail in the [Links Between Annotations](#) section.

Source mentioned in text

These are often indicated by words like *according to, ... says..., or announced by*

14. [\[Trump\]](#)<source2> It was announced today by [the U.S. Treasury](#)<source1> that additional large scale Sanctions would be **added**<event1> to those already existing Sanctions on North Korea. I have today **ordered**<event2> the withdrawal of those additional Sanctions!

15. [Trump]<source3> According to **NBC News**<source1, source2>, Voters Nationwide **Disapprove**<event1> of the so-called Mueller Investigation (46%) more than they **Approve**<event2>(41%). You mean they are finally beginning to **understand**<event3> what a disgusting Witch Hunt, led by 17 Angry Democrats, is all about!
16. [Trump] According to **Bill O'Reilly**<source1>, 80% of all the **shootings**<event1> in New York City are blacks-if you add Hispanics, that figure goes to 98%. 1% white.
17. [Trump] This is good news: @MittRomney is now **leading**<event1> in Michigan by 6 points according to **@RasmussenPoll**<source1>

Author as source

18. [Trump]<source1, source2, source3> It is my pleasure to announce that @StephenMoore, a very respected Economist, will be nominated to **serve**<event1> on the Fed Board. I have **known**<event2>Steve for a long time – and have no doubt he will be an outstanding **choice**<event3>!
19. [Trump]<source1, source2, source3> China Trade Deal (and more) in advanced **stages**<event1>. Relationship between our two Countries is very **strong**<event2>. I have therefore **agreed**<event3> to delay U.S. tariff hikes. Let's see what happens?
20. [Trump]<source1, source2, source3> Going to the White House is considered a great **honor**<event1> for a championship team. Stephen Curry is **hesitating**<event2>, therefore invitation is **withdrawn**<event3>!

Quotes

21. [Trump] "Our own Benjamin Hall is doing fantastic **reporting**<event1> on ISIS right on the from line (True). ISIS was willing to **die**<event2>but now, because of big pressure, save for a few people in caves, most have **surrendered**<event3>. A testament to our President." Thank you **Pete Hegseth @foxandfriends**<source1, source2, source3>

Retweets

Indicated by *RT* followed by the author of the original tweet, which is indicated with an '@'.

22. [Trump] RT **@WhiteHouse**<source1, source2>: The American dream is **back**<event1> because President Trump **understands**<event2> what many experts don't: that restoring American greatness i...

More than one source

23. [Trump]<source2, source3> *Every poll, Time, Drudge, Slate and others*<source1>, said I **won**<event1> both debates - but heard Megyn Kelly had her two puppets **say**<event2> bad stuff. I don't **watch**<event3>.

Confidence

Next, the annotator should score the confidence of the event with respect to the source. Note that the source is still to be used even if it is different from the author (and even when the author and source likely have differing confidence levels toward the event). Once an event has been selected, its confidence value needs to be defined as an attribute of the event annotation. This can be thought of as a score to evaluate how much the source believes in the event transpiring.

Confidence for this annotation project is scored in terms of how much the source believes that an event has taken place, is taking place, or will be taking place. For this concept, we would also like to capture positivity or negativity. Thus, we introduce an attribute for the event tag which uses a scale ranging from +2 to -2, where positive score values correspond with positive confidence and negative score values correspond with negative confidence. A score of 0 indicates complete uncertainty, with no indication in either the positive or negative direction.

For durative events, the above scale can be extended to a positive or negative belief that the event will continue to happen, so words that clearly express the cessation of durative events in the past (e.g. *stop/stopped, gave up*) will trigger high-confidence negative events.

Ex1. *I gave up listening to the teacher* (Event: *listening* → Score: -2)

For many hypothetical situations using past-tense verbs, usually introduced by *if*, scoring can be achieved by first ignoring the *if*, then assigning the event a score based on elements near it that are not hypothetical (such as *maybe, probably*), and then flipping the score.

Ex2. *If I didn't go there, I would have been fine.*

(Event #1: *go*, Non-hypothetical trigger words: *not* → Initial score: -2 → Final score: +2)

(Event #2: *fine*, Non-hypothetical trigger words: — → Initial score: +2 → Final score: -2)

Moreover, even though the event may not have happened yet, a +2 score could still be assigned if the source is confident that this event will happen sometime in the future.

Ex3. *I will eat the cake.* (Event: *eat* → Score: +2)

Ultimately, if there is great uncertainty as to what confidence score to give to the event in question, give the event a score of 0.

As stated above, the scale ranges from +2 to -2, and a more detailed break-down of these values is as follows.

Score	Name	Description	Sample Trigger Words and Phrases
+2	Positive, high confidence	The source feels sure that the event happened, is happening, or will happen (or for durative events: will continue to happen).	<i>indeed, in fact, will, must, [any positive event that is not modified by adverbs indicating uncertainty such as <i>maybe, possibly</i> or <i>probably</i>, most negative past hypothetical constructions such as <i>if s/he hadn't...</i>]</i>
+1	Positive, low confidence	The source feels there is a positive chance that the event happened, is happening, or will happen, but is not 100% sure.	<i>likely, probably, perhaps, possibly, maybe, might..., may..., could..., can...</i>
0	Neutral/uncertain	The source appears to have no confidence in either direction (positive or negative) about the happening of the event	(sentences with unknown event), <i>who knows what...?, who knows if...?, who knows how...?</i>
-1	Negative, low confidence	The source indicates that the event	<i>Unlikely, Improbably, probably ... not, might not, maybe ... not, possibly ... not, doesn't</i>
-2	Negative, high confidence	The source indicates certainty or a high confidence that the event didn't happen, isn't happening, or won't happen (or for durative events: won't continue to happen).	<i>didn't, none, no, not, doesn't, won't, failed to..., forgot to..., didn't remember to..., must not, should not, [Positive past hypothetical constructions such as, <i>if s/he had..., stopped +V, gave up +V</i></i>

Below are some examples. As in the [Events](#) section above, events are indicated in red and sourced are indicated in blue. Events that are specifically part of each category are **bolded**, whereas those that are given a different confidence score are *italicized*.

Positive, High Confidence

24. [Trump] According to **new employment numbers**<source1, source2>, 296,000 Americans have **dropped**<event1: +2> out of the work force & gave up *looking*<event2: -2> for work.
25. RT **@paulsperry_**<source1, source2, source3>: If Schiff wasn't coaching Cohen on how to **go**<event1: +2> after Trump before his testimony and was just *going*<event2: -2> over procedural issues, why didn't he invite GOP staff to *sit*<event3: -2> in on these secret meetings?

Positive, Low Confidence

26. **[Trump]**<source1, source2> Sadly, I'm probably **helping**<event1: +1> @billmaher's lowly rated show—but charity will *benefit*<event 2: +2> by \$5 million so it's worth it.
27. RT **@FAANews**<source1, source2>: Due to #HurricaneMichael, flights in your area could be **delayed**<event1: +1> and possibly **cancelled**<event2: +1>. Please continue to *check*<event2: +2> the status of...

Uncertain/Neutral

If there is no discernible event, do not annotate any of the text

28. [Trump] Merry Christmas!
29. [Trump] Despite the constant negative press covfefe
30. **[Trump]**<source1, source2, source3, source4> There hasn't been a missile or rocket *fired*<event1: -2> in 9 months in North Korea, there have been no nuclear *tests*<event2: -2> and we got back our hostages. Who knows how it will all **turn**<event3: 0> out in the end, but why isn't the Fake News *talking*<event4: -2> about these wonderful facts? Because it is FAKE NEWS!

Negative, Low Confidence

31. **[Trump]**<source1, source2, source3> @AlanDersh, a brilliant lawyer, who although a Liberal Democrat who probably didn't **vote**<event1: -1> for me, has *discussed*<event2: +2> the

Witch Hunt with great clarity and in a very positive way. He has *written*<event3: +2> a new and very important book.

Negative, High Confidence

32. **[Trump]**<source1, source2> It was a very wise move that Ted Cruz *renounced*<event1: +2> his Canadian citizenship 18 months ago. Senator John McCain is certainly no *friend*<event2: -2> of Ted!
33. **[Trump]**<source1, source2> The only thing that I have done wrong is to *win*<event1: -2> an election that was expected to be won by Crooked Hillary Clinton and the Democrats. The problem is, they forgot to *campaign*<event2: -2> in numerous states!

Emotion

Lastly, if an overt emotion is present for the author of the text, this word or words will also be annotated. It is important to note here that only emotions expressed by the author will be annotated. To be clear, if the emotion is expressed by a source other than the author, it can simply be ignored. For example, if the text said “X said that they are happy that Y”, the emotion “happy” would be attributed to X as the source of the event and not the author. Additionally, this word representing the emotion of the author should be indicating the author’s emotion towards the event or events that have been annotated in the sentence.

Once the word that represents the author’s emotion has been annotated, select an attribute that defines the type of emotion. The choices are: positive, negative, or neutral.

Positive

34. **[Trump]**<source1> I am *glad*<emotion1: positive> that my friend @EmmanuelMacron and the protestors in Paris have *agreed*<event1: +2> with the conclusion I reached two years ago.

Negative

35. **[Trump]**<source1> *Sadly*<emotion1: negative>, Democrats want to stop *paying*<event1: -2> our troops and government workers in order to give a sweetheart deal, not a fair deal, for DACA. Take care of our Military, and our Country, FIRST!

36. [Trump]<source1, source2> The Paris Agreement is fatally **flawed**<emotion1: negative> because it **raises**<event1: +2> the price of energy for responsible countries while **whitewashing**<event2: +2> some of the worst polluters....

Neutral

If there is no word or words denoting the emotion, there is no need to annotate emotion. If you find an example of an emotion word but are unsure of the author's attitude, it can be attributed a value of neutral.

37. [Trump]<source1> A vote for today's resolution by Republican Senators is a **vote**<event1: +2> for Nancy Pelosi, Crime, and the Open Border Democrats!
38. [Trump] RT @WhiteHouse<source1, source2, source3>: All across America, hardworking taxpayers **balance**<event1: +2> their household budgets, **find**<event2: +2> ways to do more with less, and still **manage**<event3: +2>...

Links Between Annotations

Once the event, source, and emotion have been tagged, links between them need to be added.

A. Emotion -> Event

A link from the emotion to the event

As seen in the example below, **Sadly** is the emotion (tagged as <emotion1>) from Trump that is linked to **paying** <event1>.

39. [Trump]<source1> **Sadly**<emotion1: negative>, Democrats want to stop **paying**<event1: -2> our troops and government workers in order to give a sweetheart deal, not a fair deal, for DACA. Take care of our Military, and our Country, FIRST!

B. Source -> Event

A link from the source to the event

As seen in example (38) below, **the U.S. Treasury** <source1> is the source for event1 and [Trump] <source2> is the source for event2. Therefore, we need to build a linkage so that we know the source is associated with an specified event. In (39), @WhiteHouse <source1, source2> is the source for both event1 and event2.

40. [Trump]<source2> It was announced today by **the U.S. Treasury** <source1> that additional large scale Sanctions would be **added**<event1: +2> to those already existing Sanctions on

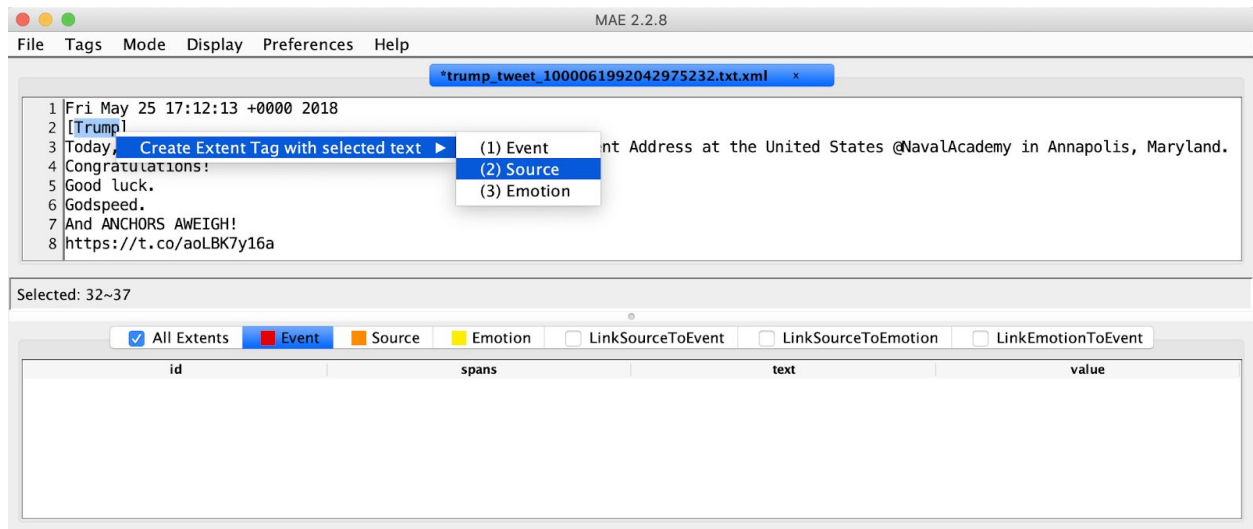
North Korea. I have today **ordered**<event2: +2> the withdrawal of those additional Sanctions!

41. RT **@WhiteHouse** <source1, source2>: The American dream is **back** <event1: +2> because President Trump **understands** <event2: +2> what many experts don't: that restoring American greatness i...

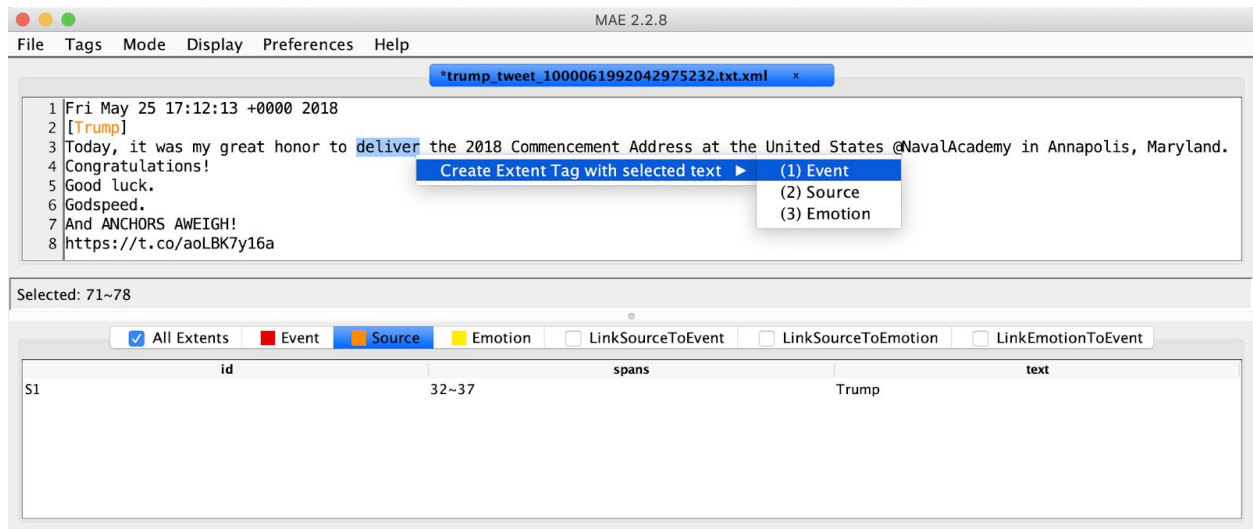
Annotation Example

[Trump]<source1> Today, it was my **great honor**<emotion1: positive> to **deliver**<source1: +2> the 2018 Commencement Address at the United States @NavalAcademy in Annapolis, Maryland. Congratulations! Good luck. Godspeed. And ANCHORS AWEIGH! <https://t.co/aoLBK7y16a>

Selecting the source:



Selecting the event:



Selecting the emotion:

MAE 2.2.8

File Tags Mode Display Preferences Help

*trump_tweet_1000061992042975232.txt.xml x

```

1 | Fri May 25 17:12:13 +0000 2018
2 | [Trump]
3 | Today, it was my great honor to deliver the 2018 Commencement Address at the United States @NavalAcademy in Annapolis, Maryland.
4 | Congratulations!
5 | Good luck.
6 | Godspeed.
7 | And ANCHORS AWEIGH!
8 | https://t.co/aoLBK7y16a

```

Create Extent Tag with selected text

- (1) Event
- (2) Source
- (3) Emotion

Selected: 56~67

☒ All Extents ☐ Event ☐ Source ☐ Emotion ☐ LinkSourceToEvent ☐ LinkSourceToEmotion ☐ LinkEmotionToEvent

	id	spans	text
S1		32~37	Trump
E1		71~78	deliver

Setting a value for the event:

MAE 2.2.8

File Tags Mode Display Preferences Help

*trump_tweet_1000061992042975232.txt.xml x

```

1 | Fri May 25 17:12:13 +0000 2018
2 | [Trump]
3 | Today, it was my great honor to deliver the 2018 Commencement Address at the United States @NavalAcademy in Annapolis, Maryland.
4 | Congratulations!
5 | Good luck.
6 | Godspeed.
7 | And ANCHORS AWEIGH!
8 | https://t.co/aoLBK7y16a

```

Selected: 56~67

☒ All Extents ☒ Event ☐ Source ☐ Emotion ☒ LinkSourceToEvent ☒ LinkSourceToEmotion ☐ LinkEmotionToEvent

	id	spans	text	value
E1		71~78	deliver	2

Setting a value for the emotion:

MAE 2.2.8

File Tags Mode Display Preferences Help

*trump_tweet_1000061992042975232.txt.xml x

```

1 | Fri May 25 17:12:13 +0000 2018
2 | [Trump]
3 | Today, it was my great honor to deliver the 2018 Commencement Address at the United States @NavalAcademy in Annapolis, Maryland.
4 | Congratulations!
5 | Good luck.
6 | Godspeed.
7 | And ANCHORS AWEIGH!
8 | https://t.co/aoLBK7y16a

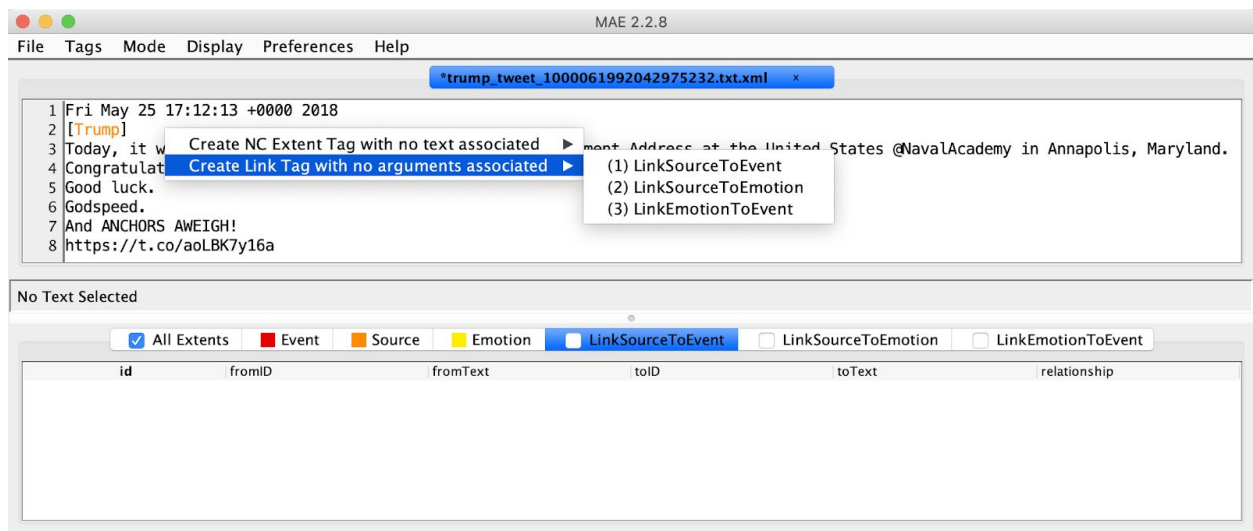
```

No Text Selected

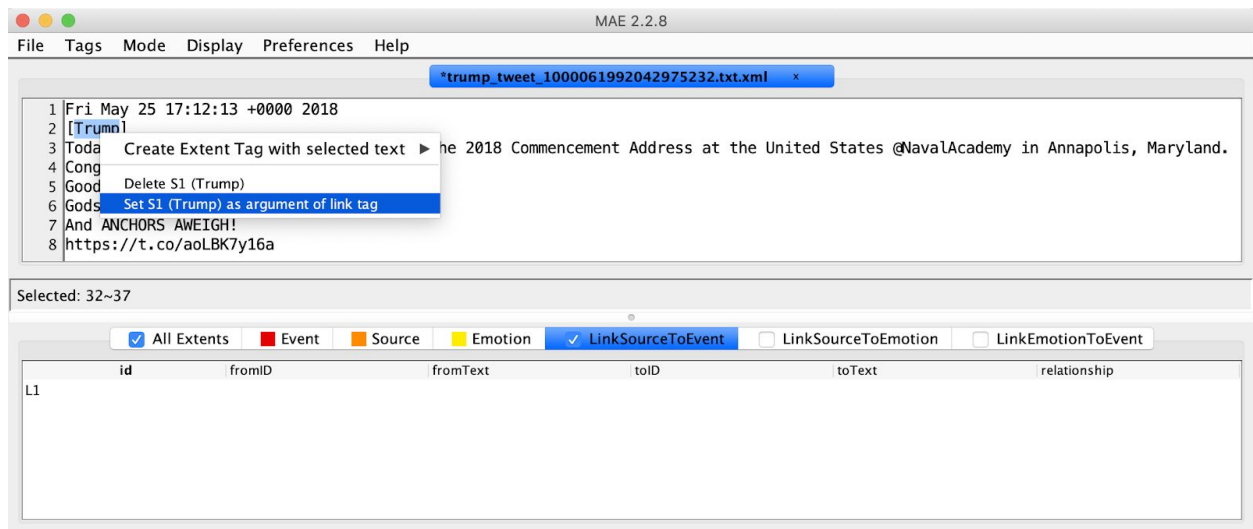
☒ All Extents ☐ Event ☐ Source ☒ Emotion ☒ LinkSourceToEvent ☒ LinkSourceToEmotion ☐ LinkEmotionToEvent

	id	spans	text	value
Em1		56~67	great honor	positive

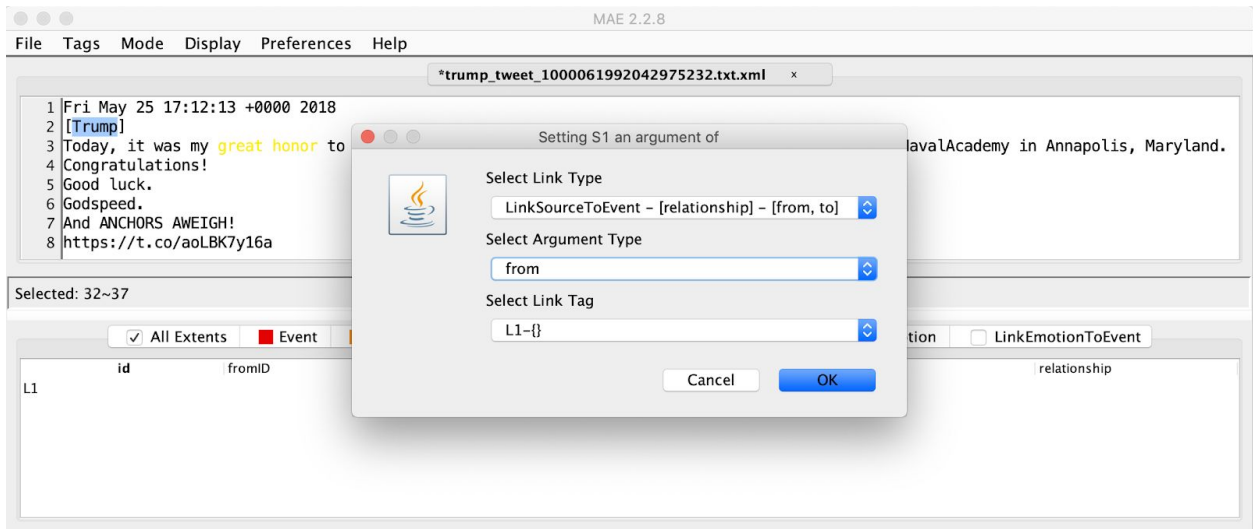
Creating a link:



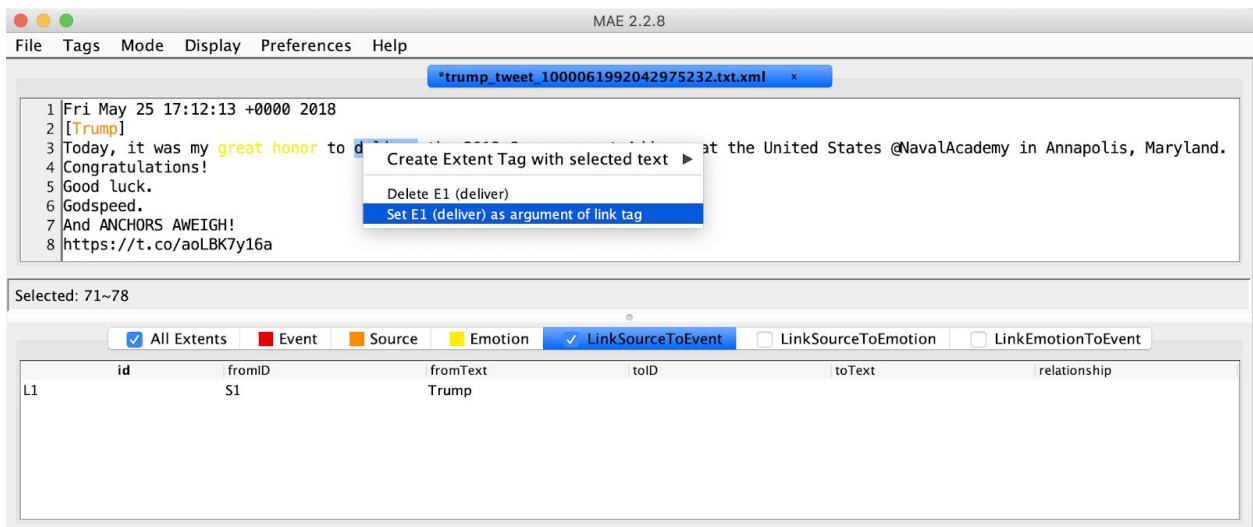
Setting first argument of the link:



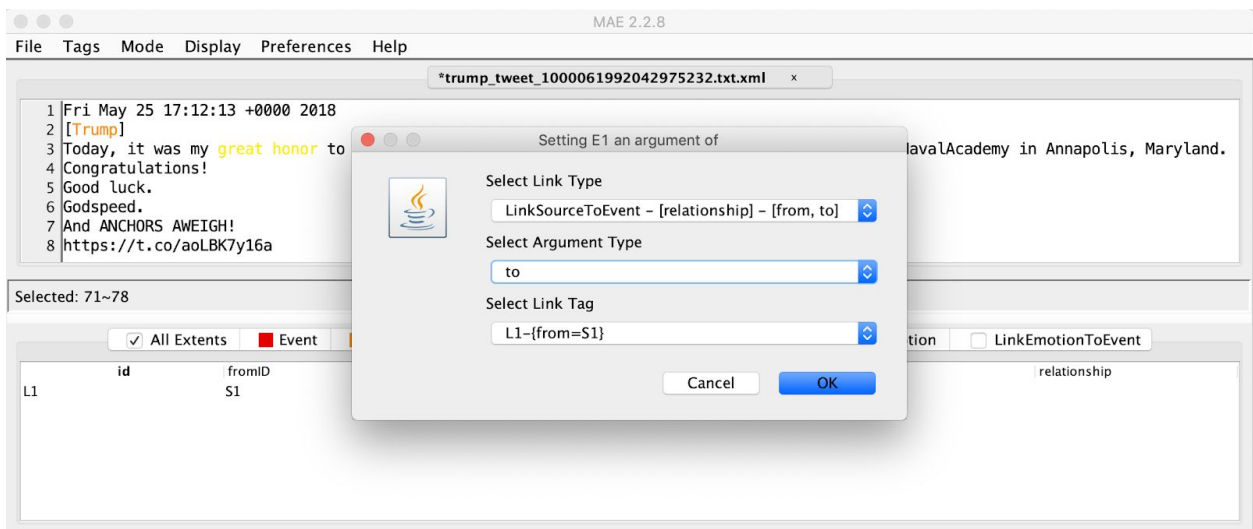
Select the desired link (when more than one) and the desired argument (from or to):



Setting second argument of the link:



Select the desired link (when more than one) and the desired argument (from or to):



Once the link is created the text will be highlighted:

The screenshot shows the MAE 2.2.8 application window. The title bar indicates the file is `*trump_tweet_1000061992042975232.txt.xml`. The main text area displays a tweet from May 25, 2018, with the text: "Today, it was my great honor to deliver the 2018 Commencement Address at the United States @NavalAcademy in Annapolis, Maryland. Congratulations! Good luck. Godspeed. And ANCHORS AWEIGH! https://t.co/aoLBK7y16a". The word "deliver" is highlighted in green. Below the text area, the "Selected: 71~78" status is shown. The interface includes a menu bar (File, Tags, Mode, Display, Preferences, Help) and a toolbar with buttons for "All Extents", "Event", "Source", "Emotion", "LinkSourceToEvent" (checked), "LinkSourceToEmotion", and "LinkEmotionToEvent". A table below the toolbar displays the following data:

	id	fromID	fromText	toID	toText	relationship
L1		S1	Trump	E1	deliver	