ANNOTATOR GUIDELINES

Thank you for helping us with this research.

You will be my *teacher* in this project. I am your student, a computer.

You will teach me to recognize *attributions* in the news.

I learn from examples that you give me.

These guidelines have been carefully written to prepare you to teach me.

They explain what an attribution is, and how to make good examples that help me learn.

Please read carefully, and be sure to let the experimenter know if something is confusing

:]

An *attribution* is basically a quotation
In this task, you'll be reading news articles and highlighting the *attributions* found within them. For the most part, *attribution* just means *quote*.
Here is an example of an attribution:

✓ Lowell said "I don't prefer pickles."

Normally, attributions occur when an author attributes some statement or attitude to another person.

The attribution doesn't have to be in quotation marks to count. Paraphrases, which describe what was said, count too, and so do mixtures of quotation and paraphrases:

✓ Lowell said that he doesn't like pickles.

✓ Lowell said that he doesn't "prefer" pickles.
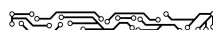
Attribution = *source + cue + content*
An attribution typically has three parts to it: a *source*, a *cue*, and some *content*:

- ✔ "The morbidity rate is striking," said *Dr. Talcott*

- ✔ Lowell said that he doesn't "prefer" pickles.

Source: the person (or thing) to which the content is attributed.

Content: the statement that is being attributed.

Cue: the "trigger phrase" (often a verb) that signals attribution is happening.

It's worth explaining the difference between the *source* and the *author*. The author is the person who wrote the articles you'll be reading, and the source is the person whom the author is quoting (or attributing to).

It's totally possible for the author to attribute *to themselves*, meaning that the source is the author, but it's not too common. Here's what that looks like

- ✔ I have said that I don't like pickles many times

Attributions include more than just quotes
An attribution occurs whenever the author attributes a *propositional attitude* to another person or thing. This includes:

- ✔ Statements made by the source
- ✔ Knowldege, belief, understanding, and contemplation by the source
- ✔ Intentions, desires, and attitudes of the source
- ✔ Source's feelings *about* something. (but not source's feelings themselves.)
- ✔ Source's support (or lack of support) for something.
- ✔ Decrees made by the source
- ✔ Facts witnessed by the source, and perceptions of the source.

We'll see examples of each of these in a moment. First, it's important to point out that attribution *doesn't* include:

- ✘ Descriptions about the source (how they feel, how they look, etc.)
- ✘ Actions done (or which will be done) by the source

Let's look at these in more detail.

✓ **Statements**

As we've seen, a statement can be attributed to a source as a direct quote, or by paraphrasing what the source said (or a mixture of the two).

✓ Lowell disclosed his relationships with the pickle regulators.

✓ "I have no regrets," said Lowell.

Note: denial counts as a statement!

✓ Lowell denies he had a role in the pickle hostage situation.

✓ Lowell denies it.

Note: negating a statement counts too!

✓ Lowell did not say that other people had been present.

✓ **Beliefs**

When a belief is attributed to a source, it's an attribution:

✓ Lowell thinks that pickles are unhealthy.

✓ The consultancy felt their lead salesperson was unfit for duty.

✓ **Intentions and desires**

Attributing intentions, commitments, or attitudes toward a thing, fact, or possibility count as attributions.

✓ Lowell plans to start his campaign for pickle health awareness.

✓ Lowell dislikes pickles.

✓ Lowell promised to double check the expiration dates.

However, attribution of intentions is tricky! If the author states that a person *will* do something, it doesn't count as attribution; but if the author indicates the person *intends* to do something, then it does. To understand why, notice that it makes sense to say "the wind *will* blow tomorrow", but not "the wind *intends* to blow tomorrow". When labelling an attribution of intention, be sure that intention really is signaled:

   ✘ Lowell will jump around.

   ✔ Lowell intends to jump around.

   ✔ Lowell plans to jump around.

   ✔ Lowell hopes to jump around.

   ✘ Lowell would jump around.

   ✘ Lowell's jump will be around.

〰〰〰

✔ **Decrees**

Reporting some decision or ruling counts as an attribution when the act of decision or ruling is based on some form of statement:

   ✔ The court ruled that Lowell's actions were lawful.

   ✔ Lowell appointed a new chief of staff.

   ✔ In July, the Environmental Protection Agency imposed a gradual ban on virtually all uses of asbestos

   ✔ The Parliament approved a ban on all uses of asbestos.

   ✔ The Court found him guilty

Think: decrees, rulings, appointments, and sometimes decisions are included because (*and when*) they involve a specific *statement*, such as a motion, a written memorandum, or an order.

〰〰〰

✓ Witnessing of facts

This case is a lot like the attribution of a statement, except that involves source noticing, remembering, contemplating, or otherwise ascertaining a state of affairs:

- ✓ Lowell knows that pickles are high in sodium.

- ✓ Lowell saw that pickles are high in sodium.

- ✓ Lowell saw a high-sodium pickle

And now some more things that seem like attribution but aren't…

✗ Descriptions about a source don't count

Descriptions about a source do not amount to statements, intentions, etc. by the source. It is easy to make the mistake of considering a statement about the emotional state of someone to be an attribution:

- ✗ Lowell was happy.

- ✗ Lowell felt happy.

- ✗ Lowell was tall.

- ✗ Lowell felt hungry because he hadn't eaten all day.

That said, if a statement, that just happens to involve emotion, is being attributed, it certainly counts:

- ✓ Lowell said he was happy.

- ✓ Lowell said "yay!" happily.

In some cases, the emotion is used to indicate the source's disposition toward a fact, which also counts:

- ✓ Lowell was happy that the ice cream was pickle-free.

- ✓ Lowell was shocked that there was so much action.

✘ **Actions don't count**

Actions don't count, so long as they aren't expressed in terms of an intention or statement of some kind:

✘ Lowell chased after the bus.

✘ Lowell will chase after the bus.

Of course, if there is a statement, intention, or desire being attributed, it counts:

✔ Lowell described how he chased after the bus.

✔ Lowell wants to chase after the bus.

✔ Lowell dreads chasing after buses.

✔ **Providing arguments for or evidence for a thing**

A source can provide evidence or arguments for some conclusion, as in:

✔ The report shows that deaths on urban interstate highways rose 7% between 1986 and 1992.

✔ His report showed that there are many risks associated with a high-sodium pickle diet.

So far so good, teacher. At this point we've covered the basics. Let's take a short break, and then try building some attribution examples.

LET'S TAKE
-A-
Coffee
BREAK
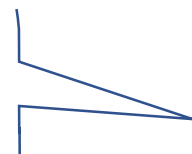
Now let's get into more detail.

Now that you've got the basic idea, let's go deeper.

As you've now seen, the task involves reading news articles, and highlighting attributions, indicating their source, cue, and content. You highlight the three parts, and then you link them together to make a nicely packaged "attribution object".

When you highlight the source, cue, and content, you need to pay close attention to the boundaries of the highlight. Be careful not to include any parts of text that don't belong (and don't miss the parts that belong!).

Pay close attention to punctuation near the beginning or end of a span that you are highlighting, and be sure to include it only when it is appropriate. We'll look closely at when punctuation should be included.

It is often difficult to decide exactly what counts as an attribution, and where exactly a source / cue / or content begins and ends. The best way to clarify that is to look at and discuss more examples together.

We'll start by taking a close look at *sources*.

Close-up on Sources

Sources can be artifacts
Sometimes the source of an attribution will be a report, a letter, a recording, a study, a database, etc. Anything that can store or relay language could itself be a source.

✔ According to the report, 53% of people are suspicious of pickles.

✔ Records indicated that he had never visited a supermarket.

The Source can be the author him/herself

If the author self-attributes, treat it just like any other attribution:

- ✓ I think he's lying.
- ✓ I hope to try again.
- ✓ There are no second chances, in my opinion.
- ✓ We just hope that they have learned something.

〜〜〜〜〜

Include modifiers and subordinate clauses that identify the source

Modifiers, like adjectives, and subordinate clauses that provide more detail about who the source is and why their commentary is interesting, should be included as part of the source:

- ✓ "Terrific," said Sterling Pratt, wine director at Schaefer's in Skokie, Ill., one of the top stores in suburban Chicago.

<div align="right">subordinate clause</div>

<div align="center">another subordinate clause</div>

- ✓ Very little is known about him, says Warren H. Strother, a university official who is researching a book on Mr. Hahn.

<div align="center">subordinate clause</div>

Not *everything* subordinate to the source should be annotated. Consider:

<div align="center">subordinate clause, but incidental</div>

- ✓ Gov. Kasich, who has the hardest "path to victory" for his party's nomination, said last night: "We'll have another rapprochement with Turkey."

Here the subordinate clause "who has the…nomination" seems *incidental*, rather than being an *important part of identifying the source*, so we exclude it. To decide if elaborating text should be included in the cue, ask yourself these questions:

(1) Would it help you identify the source?
(2) Does it establish the relevance of the attribution?
(3) Does it help you interpret the statement, e.g. by establishing credibility?

Now consider the following example, where the annotator chose to extend the source to include a subclause:

✗ Jeremy, who supports Sanders, says he'll definitely vote tomorrow.

As indicated, here the source should not really be extended. Notice that, extending the source actually masks an attribution; the correct annotation indicates two attributions:

✓ Jeremy, who supports Sanders, says he'll definitely vote tomorrow.
✓ Jeremy, who supports Sanders, says he'll definitely vote tomorrow.

Include determiners, quantifiers, and adjectives:
Whenever a span has a determiner ("*the* reporter", "*an* eye-witness") or a quantifier ("*some* people", "*most* republicans"), or an adjective ("*diligent professionals*") the determiner / quantifier / adjective should be included in the same span. This generally applies to the source, cue, and content, though is seen far more often for the source. (Although we will see later that there is an exception for cues, due to the single-contiguous-cue rule.)

Here is a somewhat less straightforward example (it lacks an explicit source, which will be covered later) that shows this rule in action:

✓ The so-called "Human Lizard" is sure to attract attention.
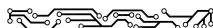
Sometimes sources are possessive pronouns
Sometimes we have a case where the source is a possessive pronoun, and the cue is a noun signifying an act of communication:

✔ Her words: "Don't panic".

✔ Her advice was to "stay calm".

This typically happens when the cue takes the form of a noun that signifies an act of speech or other expression.

It is common, in such cases that there is some tension between the source and cue: some words may seem to be part of both the source and the cue.

We'll discuss such cases in detail, when we talk more about the cue momentarily. For now, consider this example:

✘ The president's report indicated that there was nothing to panic about.

Although the annotation above seems reasonable, the correct annotation is the one below:

✔ The president's report indicated that there was nothing to panic about.

This example illustrates that there can be tension between the source and the cue. The first annotation above takes the opinion that the artifact produced by the president is the source (and "The president's" is included as a modifier). The second annotation takes the opinion that the source is the president, and the artifact is part of the process of expression. The ambiguity arises because an artifact can be seen as a source or as part of the expression process. In such cases, we take the creator of the artifact to be the source, and the artifact to be part of the cue. The artifact generally only occupies the source role when its creator is out of scope, and so the artifact becomes a character in the news story.

<hr/>

<u>Attributions can share sources</u>
It's possible for two attributions to share the same source. Consider the following excerpt:

- Newsweek, trying to keep pace with rival Time magazine, announced new advertising rates for 1990 and said it will introduce a new incentive plan.

This example has *two* attributions in it, which both have "Newsweek" as

their source:

✓ Newsweek, trying to keep pace with rival Time magazine, announced new advertising rates for 1990 and said it will introduce a new incentive plan.

✓ Newsweek, trying to keep pace with rival Time magazine, announced new advertising rates for 1990 and said it will introduce a new incentive plan.

Some attributions have compound sources

The source may be made up of multiple agents, as in this example:

✓ Google and Yahoo say they are prepared to take the issue to court.

Include "and"

Notice that we include the conjunction ("and"), because we consider the whole noun-phrase to be the source.

Some attributions have *no* source

Attributions can sometimes have no source at all. This happens when the source of the attribution is implicit, which tends to happen when the attribution is done in a passive voice:

✓ It is said that no one has ever returned from such a bad streak.

✓ She was known as the rainmaker.

Note:
An attribution has:
- *zero* sources, *one* source, or *a compound* source,
- *one* cue,
- at least *some* content.

Beware passive constructions

As we just saw, passive constructions enable attribution without an explicit source. But sometimes, the source is found to be a prepositional phrase dependent on a verbal cue:

✔ It has been said by many that one should buy low and sell high.

✔ There is concern among his colleagues that they will eliminate the position.

Empty pronouns in passive constructions are not sources

The second excerpt from the previous above illustrates another useful rule: the pronouns "it" and "there" are often introduced in forming passive constructions to satisfy the grammatical requirement for a subject, but are empty. They should not be annotated:

Not a valid source!

✘ There is concern that they will eliminate the position.

✔ There is concern that they will eliminate the position.

✔ It is said that there will be rain tomorrow.

Well done for making it this far, teacher!

As you've seen, the source of an attribution can come in many shapes and sizes.

And sometimes is not even there.

So much for sources. It's time to take a deep dive into *cues*. Forge ahead!

Close-up on Cues

A variety of words and phrases can serve as cues

Although the most common cue is "said", a variety of words can serve as cues:

  ✓ knows, showed, write, rejected, ruled, plans, felt, thinks, confirms

Sometimes, a manner verb is used to connect the source to the content, thus indicating attribution:

  ✓ "They'll pay for what they did," smiled Lowell.

Here, "smiled" essentially stands for "said while smiling".

<hr/>

Nouns as cues
Usually the cue is a verb, but sometimes it is a noun. Consider this example:

  ✓ The volatility isn't entirely bad, in Mr. Connolly's view.

"view" is what connects Mr. Connolly to the content and lets us know that attribution is happening, therefore, it is the cue. Often, as we see here, when the cue is a noun, the source is the possessor.

Frequently, when the cue centers on a noun, it still involves a verb to connect it to the source or content:

  ✓ His concern is that they will eliminate the position.

                    verb

  ✓ He raised the concern that they will eliminate the position

<hr/>

Prepositions and prepositional phrases as cues
Cues can also be *prepositions and prepositional phrases*, for example:

  ✓ There was a story about reconstruction era politics.

  ✓ According to the Encyclopedia Britannica, Earth has seven continents.

  ✓ The bailouts are unjust, in the eyes of most Americans.

Remember that the function of the cue is to signal attribution; therefore,

the cue should comprise a phrase that is sufficient to do that job.

The single preposition "about" can do that on its own, but prepositions like "of", "to", etc. are (with some exceptions) usually insufficient, and need to combine with other words to signal attribution. For example:

✗ The statement of defiance was not well-received.

✓ The statement of defiance was not well-received.

Here, the word "statement" is needed in the cue. In such a case, the source is left as implicit. This means we think of the last example as being a paraphrase of this:

✓ She made a statement of defiance that was not well-received.

〰〰〰

"That" is not a preposition!
Be aware that the word "that" is *not* a preposition. Often it is used as a conjunction to introduce content, and is part of the content:

✓ Most Americans believe that the bailouts are unjust.

include in content

Although "To" is a preposition, it is often part of the content!
When the word "to" appears at the boundary between a cue and content, it is usually part of the content. That's because often it is part of the infinitive form of the first verb in the content:

✓ She hopes to become the president.

✗ She hopes to become the president.

〰〰〰

The cue should indicate attribution, not just action.
It can be difficult to decide what is part of the cue. Often, some words seem like they might be part of the content, but really, they should be included in the cue. Let's look at an example where that happens:

✗ The group is attempting to highlight that some are struggling.

✘ The group is attempting to highlight that some are struggling.

✔ The group is attempting to highlight that some are struggling.

Why are the first two annotations wrong? The main reason is that neither "is" nor "is attempting to" are valid cues—they don't signal attribution. (Recall that *attempting something* is considered an action, but not an attribution.)

So, sometimes words seem like they should be part of the content, but really they should be part of the cue.

In a similar way, sometimes words seem like they should be part of the source, but are really part of the cue.

✘ She made a statement of defiance, which was not well-received.

✔ She made a statement of defiance, which was not well-received.

When deciding what is the cue, always start out by identifying the minimum set of words that really indicates that an attribution is occurring. The word "of" just isn't specific enough to indicate attribution on its own. In the example above, the word "statement" is what signals attribution. Here "statement" isn't the source, it refers to the act of expressing that is being attributed to the true source, "she".

We've already seen the next example, but it shows how to handle a case where the source and the cue seem to both compete for the same word:

✘ His concern is that they will eliminate the position before he returns, which is understandable.

✔ His concern is that they will eliminate the position before he returns, which is understandable.

So, what's wrong with the first annotation above? The word "is" doesn't specifically indicate attribution, so we need to include the word "concern". In this case, the source is indicated by the possessive pronoun "his".

A general principle to apply, which is consistent with these examples, is that the source should be the sentient agent (if one is available), and the cue should encompass the expressive artifact, act of expression, or the

holding of an internal state.

Keep the verb group together

As we've already seen, cues often contain verbs. Whenever that is the case, you should be sure to keep the "verb group" together.

What is a *verb group*? Verbs can be accompanied by an entourage, including adverbs, auxiliary verbs, and prepositions. In this task, we're calling the verb together with its entourage, the "verb group". Whenever a verb is included in a cue, its entourage should be included too! So, this includes:

✔ *auxiliary verbs*: has, have, had, be, is was, are, do, does, did, …

✔ *modals*: should, could, would, might, may, ought, will, …

✔ *negation*: not, won't, didn't, wasn't, isn't, hasn't, …

✔ *adverbs*: clearly, loudly, softly, aggressively, quickly, …

✔ *prepositions*: on, in, at, for, about, upon, to, under …

Have a look at these examples, and notice which words are included in the cue:

✔ She could not say whether she had seen the assailant earlier in the day.

✔ They blurted immediately something to the effect of there having been foul play.

✔ The product release was probably announced too early.

✔ They have repeatedly called for a cease fire.

Preposition linking cue to content

✔ He should not have regularly talked about the incident.

✔ "You'll pay for this!" she said with anger.

✔ "You'll pay for this!" she said angrily.

Here is a more complex example that illustrates (a) keeping the verb group

together and (b) a passive attribution:

- ✓ The news lead to calls by some Democrats for the Obama administration to blame Russia publicly.

- ✓ The news led to calls by means of telegraph for the Obama administration to blame Russia publicly.

In both cases, the core of the cue is the word "calls". Now, in the second case, we have replaced the source with a modifier which should be part of the cue.

Only one contiguous cue span per attribution
You will use this rule often, so we will name it the *single-contiguous-cue* rule. You may notice that in the first example, the word "for" is not annotated as part of the cue, whereas it is part of the cue in the second example. One might suggest that it should be annotated in both cases, given that they are both modelled after this kind of attribution, in which "for" is part of the cue:

- ✓ Some Democrats called for the Obama administration to blame Russia publicly.

However, in this task, we will adopt the strict rule that there is only one uninterrupted cue span allowed per attribution. In the first example, we can't include "for" because it is detached from the rest of the cue span by the source. In such cases where the cue has been split up, we choose the span which most explicitly signals attribution.

Punctuation as a cue
You may have noticed previous examples where a colon (":") was marked as part of the cue. Sometimes punctuation is used to indicate that attribution is happening (and hence it is part of the cue). When speech is introduced with a colon, the colon is part of the cue:

- ✓ He said: "Don't make me come over there!"

- ✓ Romeo: "Is the day so young?"

- ✓ Her advice: "Don't panic."

Note that quotation marks are considered part of the content.

Another example where punctuation acts as a cue is when a quote is tagged afterward, using a dash (—) followed by the speaker's name:

✔ Imagination is more important than knowledge.
— Albert Einstein

A comma at the end of a cue should usually **not** be included. There are exceptions, and we'll cover this in more detail in the section on punctuation, but here is an example that illustrates a frequent pattern:

don't include!

✔ As one source said, "This is yet another step to replace public institutions with private businesses."

Cues can be compound
Occasionally you may come across a compound cue. In that case highlight the full compound cue, including the conjunction:

✔ He said and believes that the Earth is flat.

include "and" (the conjunction)

Choose the sentient source

In cases where the rules mentioned so far leave open multiple possibilities for annotating the cue, you should generally favor the annotation that selects a language-capable agent as the source, even when that relegates an artifact as the cue.

Consider these alternative annotations:

✔ The group also pasted up posters explaining their philosophy.

✘ The group also pasted up posters explaining their philosophy.

✔ She issued a statement indicating that violence won't be tolerated.

✘ She issued a statement indicating that violence won't be tolerated.
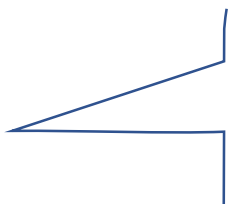
Phew, that was hard!

But the hardest part is behind you teacher!

The next bit easier.

Once you finish it, we'll take a break.

Closeup on Content

Be conservative, only include what is clearly attributed
The content consists of everything that the author attributes to the source.
Sometimes it may be unclear what is being attributed to the source, and
what is merely added by the author. Consider this excerpt:

> ✓ They announced plans for three more facilities, which will likely
>   be located in Europe due to its favorable regulations.

In the above example, we can't be sure whether the second clause ("which
will likely be located …") was part of the announcement. It seems like it
could be background information added by the author. When in doubt,
don't include it in the content.

Don't include the recipient of the statement
Sometimes an attribution will involve the source making a statement *to a
recipient*. In general, the recipient should not be highlighted:

> ✓ She is telling me otherwise.

don't highlight the recipient

Here "me" is not part of what is being said. It's *who* it is being said *to*.

But, don't confuse the recipient with a person who is rightfully part of the
content:

> ✓ She said he doesn't dance.
> ✓ She calls him the man with two left feet.

In these examples "he" and "him" are really part of *what* is being said.

Content continuations should be included in the attribution

A common pattern is to stick the cue and source between two pieces of content. Both pieces of content should be annotated as part of the same attribution:

- ✔ "We'll be here for a while yet," she said. "These cases take a lot of patience."

- ✔ "There's no question that some of those workers and managers contracted asbestos-related diseases," said Darrell Phillips, vice president of human resources for Hollingsworth & Vose. "But you have to recognize that these events took place 35 years ago"

Notice that the continued content is part of the same attribution, even though it starts a new sentence.

Sometimes the continuation will appear in its own paragraph. It should still be annotated as part of the same attribution:

- ✔ "Before long, we'll run out of personnel to cover the region," said Mr. Dunlap, Ridgeford County's Police Chief.

  "I warned them years ago, but nothing. It's irresponsible. That's what it is—irresponsible."

Occasionally, you might see the isolated, quoted paragraph appearing *before* the paragraph that contains the cue and source.

When trying to divide content between attributions, remember that *every attribution has just one cue*. Deciding what attribution content belongs to is about deciding which cue it is bound to.

Notice how this differs from attributions sharing the same source:
- ✔ Newsweek, trying to keep pace with rival Time magazine, announced new advertising rates for 1990 and said it will introduce a new incentive plan.

- ✔ Newsweek, trying to keep pace with rival Time magazine,

announced new advertising rates for 1990 and said it will introduce a new incentive plan.

In that case, we create two separate attributions because there are two distinct cues: "announced" and "said".

<div align="center">⌇⌇⌇⌇⌇⌇</div>

Closeup on punctuation

Include punctuation when it is part of the span
A basic rule for whether to include punctuation in a source, cue, or content span is to include it if it is fully within the span. When the punctuation is at the beginning or ending of the span, it should normally be excluded, but there are exceptions.

Let's look at examples showing how to handle punctuation, starting with a look at punctuation relating to sources.

<div align="center">⌇⌇⌇⌇⌇⌇</div>

Include commas at the end of a span when part of a parenthetical pair
Recall that, when there's a relative clause that elaborates and helps identify a source, we include it as part of the source. Often, the subordinate clause will be parenthetical, with commas that act like parentheses. Parenthetical commas are part of the source:

included

Kerry, who lost the presidency to Bush in 2004, said he was a "real gentleman, who would disagree with that?"

included

<div align="center">⌇⌇⌇⌇⌇⌇</div>

Handling punctuation in the content
We should only include punctuation in the content when it is really part of the content. Consider the excerpt below. We have not included the period at the end of the sentence, because it is ending the whole sentence and is not actually part of the content.

✔ John, still breathing heavily, said he would continue his daily training regimen.

exclude

21

Compare that to the next excerpt. The question mark is part of the content that is being attributed, so it should be labelled as such:

include

✓ She wondered, would there be any retribution?

In the content—include everything that is within quotes

This is fairly straightforward:

✓ "I haven't heard anything," she said, "but that doesn't mean it isn't true."

exclude

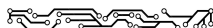include

include

Ok Let's take a short break!

Welcome back!

You're starting to become an expert attribution annotator! You're almost ready to start teaching me to be just as good as you.

In your work so far, you may have come across some tricky cases.

Let's have a look at how to handle the really tough examples!

Special cases and grey areas

No nested attributions

It is common to see an attribution *within the content* of another attribution. Consider these two alternative annotations of an excerpt:

✘ Reardon, head of the workers' union, criticized Mr. Lamport for saying that a minimum wage is bad for business.

✔ Reardon, head of the workers' union, criticized Mr. Lamport for saying that a minimum wage is bad for business.

As you can see, the excerpt above has two potential attributions—one having Reardon as the source, the other having Mr. Lamport. But we only annotate the outermost attribution (the one with Reardon as the source), and ignore the nested attribution to Mr. Lamport.

The excerpt below also has a nested attribution, can you see it?

✔ Attempts to hack into Republican political organizations over an extended period were reported in intelligence bulletins circulated by U.S. agencies, four sources said.

No shudder quotes or title quotes

Quotation marks can be used for purposes other than quoting speech, e.g. to distance the author from a particular phrasing (called "shudder quotes"):

✘ But the "terrorists" were women and children who were trying to escape the harsh living conditions of their home country.

✘ It wasn't so much a "truce" as it was mutual exhaustion.

Quotes can also be used to isolate a name, title, or label:

✘ Later he appeared on "Meet the Press."

✘ The examples above illustrate "shudder quotes."

Such quotation marks do not signal attribution!

～～～～

Titles of works

A potential pitfall occurs with references to titles of a written or artistic work. In certain cases, expressing a work's title can be structured like an attribution:

✔ The collection is called Jabba the Hutt: The Art of the Deal

✘ The collection "Jabba the Hutt: The Art of the Deal" is popular.

✔ The poem is known as "Ode to Ye Olde."

✘ The title of the poem is "Ode to Ye Olde."

～～～～

Attribution of work

This task is all about the attribution of *propositional attitudes*: statements, beliefs, intentions, desires, decrees, the witnessing of facts, etc. This does *not* include the attribution of work, like books, plays, paintings, etc. So, neither of these should be annotated:

✘ Alfred Hitchcock wrote Haunted Houseful.

✘ Alfred Hitchcock wrote a book.

On the other hand, the following excerpts *do* include attribution. Note that in these examples, actual content of the work is being attributed:

- ✔ Alfred Hitchcock wrote a story about ghosts and ghouls.

- ✔ Alfred Hitchcock wrote "Haunted houseful", a story about ghosts and ghouls.

- ✔ Alfred Hitchcock wrote the story "Haunted Houseful" about ghosts and ghouls.

Attribution to visual media

A visual medium can be the source for an attribution. The essential test is whether there is a cue phrase that indicates an act of expression or representation:

- ✘ It was a photo of a man with a balloon.

- ✔ It was a photo depicting a man with a balloon.

- ✘ The movie combines modern cinematography with the pacing of a play written for the stage.

It is worth looking closely at the examples above: although they are essentially the same, the first does not amount to an attribution because of how it is phrased. The third is description of the movie, but does not attribute content.
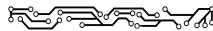
Indirect content

In some cases you will see an indirect reference to content. Consider the excerpts below:

- ✘ John only said three words.

- ✔ She denied it.

- ✘ John made a statement.

We don't consider the first attribution to be an attribution, because "three words" refers to the *statement* but not the *content*. I.e. "three words" refers to the physical manifestation of the statement but not what it means.

On the other hand, we consider the second excerpt to be an attribution, because here "it" stands for the proposition that she is denying. If she is denying the allegation that she participated in a burglary, then "it" stands for participation in a burglary:

✔ She denied it.

✔ She denied participating in the burglary.

"Try" as a cue

When you see something that looks like an attribution with "try" as the cue, it might not be an attribution:

✘ He was trying to stop her.

✘ They will try to release the product on schedule.

Try signifies the concrete action of attempting, and doesn't specifically signify intention (even though it does imply an intention).

Although "try" on its own doesn't generate an attribution, it could still occur within a cue involving other elements:

✔ She plans to try it.   ← intention

✔ He tried to tell her not to.   ← statement

✔ He was trying to convince her to stop.   ← statement

Although actions imply an underlying intention, we should always focus on what the text *actually says*. If the text attributes an intention, that's an attribution; if the text describes an action (including an attempt) it's not an attribution.

"Seek" can indicate an action *or* an intention / desire

When you see the word "seek" as a potential cue, ask yourself whether it refers to the act of searching, or to a desire or intention. To help you differentiate these possibilities, consider the examples below:

✘ She's seeking clues to the crime. ← action ("seek" = "search for")

✔ She only seeks fame and fortune. ← desire ("seek" = "wants")

Negated attribution still counts!
If the act of attribution is itself negated, you should still annotate it in the usual way. Note that there still must be attributed content!

✔ They did not disclose the identities of the organizations.

✔ Lowell did not say "I like pickles."

Declining and refusing are acts of expression

To declining or refuse does not mean to simply avoid an action, but to express the intention not to do that thing. Therefore, we annotate such cases:

✔ He declined the invitation to go camping.
✔ He declined to go camping.
✔ He declined to comment.
✔ He refused to comment.
✔ He said "No comment."

Notice the difference with simply not doing something.

✘ He did not go camping.
✘ He did not comment.

You may be tempted by the word "comment" in the last example, however, it is an act of expression, not content. It can be part of a cue, but there must be some content as well:

✘ He did not comment about his alleged involvement.

Generally, include the conjunction when a span is compound

- ✓ They were charged with "disturbing the peace" and "indecent exposure."

- ✓ Yahoo and Google said they will take the matter to court.

- ✓ He said and believes that the Earth is flat.

Relative pronouns

We already saw that possessive pronouns can be sources. Now, we turn to cases where the source is referred to via a *relative* pronoun like "which" or "who". In these cases, it can be tempting to "resolve coreference" and include both the relative pronoun and its coreferent mention. But don't do that:

relative pronoun

- ✓ Bay Financial, which said it may be forced to file under Chapter 11, plunged 1 3/8 to 2 1/8.

- ✓ The Wall Street trader, who claimed to have no insider knowledge, is likely to evade charges.

relative pronoun

Notice that "Bay Financial" and "The Wall Street trader" are left out of the source span, because including them involves coreference resolution.

A similar phenomenon can take place with content:
- ✓ He was the man who the judge accused of breaking the rules.

Here "the man" is left out of the content span because it would involve coreference resolution for "who".

Wow, you're all done!

You're officially an attribution annotation expert!

I can't wait to get started learning from you, teacher!

28