

# Annotation Guidelines for the **REFreSD** dataset

---

- ✓ **Dataset name:** Rationalized-English French Semantic Divergences
  - ✓ **Citation:** Detecting Fine-Grained Cross-Lingual Semantic Divergences without Supervision by Learning to Rank, *In EMNLP 2020*
  - ✓ **Data developers:** Eleftheria Briakou & Marine Carpuat
  - ✓ **Link to dataset:** <https://github.com/Elbria/xling-SemDiv/tree/master/REFreSD>
- 

You are asked to compare the meaning of English and French text excerpts. You will be presented with one pair of texts at a time (about a sentence in English, and a sentence in French). For each pair, you are asked to do the following:

1. Read the two sentences **carefully**. Since the sentences are provided out of context, your understanding of content should only rely on the information available in the sentences. There is no need to guess what additional information might be available in the documents the excerpts come from.
2. Highlight the text spans that convey different meaning in the two sentences.

## Guidelines on highlighting spans:

- After highlighting a span of text, you will be asked to further characterize it as:
    - **Added** Choose this class if the highlighted span corresponds to a piece of information that does not exist in the other sentence.
    - **Changed** Choose this class if the highlighted span corresponds to a piece of information that exists in the other sentence but their meaning is not the exact same.
    - **Other** Choose this class if none of the above holds.
  - You can highlight as many spans as needed.
  - You can **optionally** provide an explanation for your assessment in the text form under the Notes section (e.g., literal translation of idiom)
3. Compare the meaning of the two sentences by picking one of the three classes
    - a) **Unrelated** The two sentences are completely unrelated or have a few words in common but convey unrelated information about them
    - b) **Some meaning differences** The two sentences convey **mostly the same information, except for some details or nuances** (e.g., some information is added and/or missing on either or both sides; some English words have a more general or specific translation in French)
    - c) **No meaning difference** The two sentences have the **exact same meaning**

## Task Interface

Below we include screenshots of the task interface at each of the steps (1,2,3) described in the *task guidelines* section. We use an example sentence-pair.

1.

English-French annotation task

Example: 1/ 59.

Highlight the text spans that are different in meaning between the two sentences:

In 1853, Toyo was appointed by the head of Tosa domain Yamanouchi Toyoshige to reform and modernize the domain.

En 1853, il est nommé à la tête du domaine de Tosa par son daimyo Yamanouchi Toyoshige pour le réformer et le moderniser.

[Label Me](#)

Assess the relation of the two sentences.

2.

English-French annotation task

Example: 1/ 59.

Highlight the text spans that are different in meaning between the two sentences:

In 1853, **Added** Toyo was appointed by **Changed** the head of Tosa domain Yamanouchi Toyoshige to reform and modernize the domain.

En 1853, il est nommé à la tête du domaine de Tosa par son **Changed** daimyo Yamanouchi Toyoshige pour le réformer et le moderniser.

[Label Me](#)

Assess the relation of the two sentences.

**Possible class name of highlighted text:**

- Added
- Changed
- Other

### 3.

English-French annotation task

Example: 1/ 59.

Highlight the text spans that are different in meaning between the two sentences:

In 1853, <sup>Added</sup> **Toyo** was appointed by <sup>Changed</sup> **the head** of Tosa domain Yamanouchi Toyoshige to reform and modernize the domain.

En 1853, il est nommé à la tête du domaine de Tosa par son <sup>Changed</sup> **daimyo** Yamanouchi Toyoshige pour le réformer et le moderniser.

**Label Me[Some\_meaning\_differences]**  
Assess the relation of the two sentences.

Possible class name of sentence-pair:

Label Me:

**Notes**

- Unrelated
- Some\_meaning\_differences
- No\_meaning\_difference