Recruiting email for the REFreSD dataset

- ✓ **Dataset name:** **R**ationalized-**E**nglish **Fre**nch **S**emantic **D**ivergences
- ✓ **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**

Want to earn a $30-$267 Amazon Gift card?

We are running an annotation task to collect human judgments on the meaning differences observed in English-French text excerpts. Participants should fulfill the following conditions:

a) UMD students or community members at least the age of 18
b) proficient in **both English and French** languages
c) background in Languages and/or Translation studies preferred (not required)

Following we include further information on the annotation procedures and logistics.

Annotation Procedures

Annotators will be asked to complete the task using a web interface on one of the provided laptops. In each session, a participant will be presented with at most 120 instances to annotate. Each instance consists of (a) a short text excerpt in English, and (b) a short text excerpt in French. Annotators are first asked to highlight text spans that differ in meaning across the two excerpts. Then, they are asked to assess whether the texts as a whole are unrelated, have the exact same meaning, or fall somewhere in between.

Logistics

- ■ **Duration:** The annotation will be done in sessions of about 2 hours each. During each session, you will be presented with about 120 examples to annotate. While you are encouraged to participate in up to 8 sessions you can stop after a single session or anytime during a session.
- ■ **Place:** The study will take place either online (you can use your own computer as long as it is connected to a UMD network e.g., via VPN).
- ■ **Compensation:** Annotators will be compensated at a rate of $2 per 10 instances, with a bonus of $5 for completing the first session of 120, and $10 for completing additional sessions after the first.

Who are we?

- ✓ Researchers from the Department of Computer Science at UMD
- ✓ Members of the Computational Linguistics and Information Processing lab