Ethical Review for the REFreSD dataset

- ✔ **Dataset name:** **R**ationalized-**E**nglish **Fr**ench **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
- ✔ **Review Board:** University of Maryland College Park Institutional Review Board
- ✔ **Revew Type:** Expedited Review
- ✔ **Link to dataset:** **https://github.com/Elbria/xling-SemDiv/tree/master/REFreSD**

PROCEDURES

The procedure involves completing an annotation task on a laptop. 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) or in person at the Brendan Iribe Center for Computer Science and Engineering on the UMD campus. The annotation will be done in sessions of about 2 hours each. While you are encouraged to participate in up to 8 sessions, you can stop after a single session or anytime during a session. During each session, you will be presented with about 120 examples of short text excerpts in French and in English (about a sentence long). For each example, you will read the text in both languages and answer simple questions about the differences in meaning between the content conveyed in each language.

POTENTIAL RISKS

There are no known risks beyond that of using a computer associated to participants of this study. If participants become tired and/or bored during the annotation process they are free to take a break or stop anytime.

POTENITAL BENEFITS

There are no direct benefits from this study. The potential benefits to you include the ability to better study how and when the translation process introduces meaning divergences across languages. We hope that, in the future, other people might benefit from this study through improved understanding of semantic divergences in bilingual texts, which can help improve Natural Language Processing technology, and design better tools to support second language learners, translators or web users interested in accessing content in many languages.

CONFIDENTIALITY

We will not ask you for any personal information beyond your email address. You can sign this consent form electronically and return it to the investigators via email. Any potential loss of confidentiality will be minimized by storing data and consent forms securely in UMD-controlled cloud storage (box). Before the experiment starts, you will be assigned a User ID by the investigators. This is a random number that is not tied to any sensitive nor personal information. Only the investigators will have access to the mapping between (a) your User ID and (b) your name and email. Upon completion of the experiment, data will be anonymized and released publicly. If we write a report or article about this research project, your identity will be protected to the maximum extent possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you or someone else is in danger or if we are required to do so by law.

COMPENSATION

You will receive Amazon gift cards for an amount determined by the number of completed annotations, on the basis of $2 per batch of 10 examples, with a bonus of $5 for completing the first session, and $10 for completing additional sessions. The maximum amount of money a participant can get for completing all 8 sessions is $267. You will be responsible for any taxes assessed on the compensation.

CONSENT PROCESS

All participants will be presented with informed consent and asked to sign a consent form electronically before the start of the annotation process. Consent forms will be written in English. All participants will receive a copy of the consent form for their records. If the session takes place in person, participants will sign the informed consent form electronically at Brendan Iribe Center for Computer Science and Engineering, behind closed doors, to protect their privacy. If the session takes place remotely, participants will sign the consent form from their homes and return it to the investigators via email in a pdf format.

RIGHT TO WITHDRAW

Your participation in this research is completely voluntary. You may choose not to take part at all. If you decide to participate in this research, you may stop participating at any time. If you decide not to participate in this study or if you stop participating at any time, you will not be penalized or lose any benefits to which you otherwise qualify. **If you are an employee or student, your employment status or academic standing at UMD will not be positively or negatively affected by your participation or non-participation in this study.** If you decide to stop taking part in the study, if you have questions, concerns, or complaints, or if you need to report an injury related to the research, please contact the investigator.