Hypotheses Study 1

The first hypothesis concerns the average absolute sentiment intensity of conjoined adjectives in each corpus: we expect the legal corpus (LC) to contain conjunctions with more neutral sentiment values than the Reddit corpus (RC) (H_1).

Secondly, we expect the legal corpus to have more neutral conjoined sentiment values on average, and that the effect is significant for both positive and negative target adjectives (H_2). This results in the following directed hypotheses: On the one hand, H_{2a} posits that conjoined sentiment values for negative target adjectives are on average significantly closer to 0 for LC compared to RC. H_{2b}, on the other hand, assumes the same for positive target adjectives. **Note:** In contrast to H_1, H_{2a-b} are not only interested in sentiment intensity, but also also in polarity. Thus the second model uses the non-transformed sentiment values as dependent variable, instead of the absolute sentiment values used to test H_1.

The third hypothesis is that the difference between the corpora persist across all concept classes. We expect more neutral average sentiment values for each concept class in LC compared to the same concept class in RC (H_3). The partial hypotheses are defined as follows: H_{3a} assumes that epistemic concepts have lower average conjoined sentiment values in LC compared to RC. H_{3b} expects that the same holds for legal concepts, while H_{3c} does so for TC. **Note:** Because we are mostly interested in differences in terms of intensity, the third model once again uses absolute conjoined sentiment values, as for testing for H_1. As estimator we use an interaction term between the corpus dummy (RC/LC) and the concept class factor (Epistemic/Legal/TC).

Hypothesis Study 2

In the second study, we focus on LC only. In order to be able to compare the results of the evaluative concepts classes to a baseline, we added corpus entries for the following descriptive target adjectives: *active*, *ambiguous*, *complex*, *explicit*, *limited*, and *practical*. Instead of comparing context effects (Study 1), we want to inquire whether the concept classes (Descriptive/Epistemic/Legal/TC) can be distinguished from each other within the legal context. H_4 posits that the four concept classes have significantly different sentiment averages for their respective adjective conjunctions. **Note:** The linear model includes the absolute conjoined sentiment as dependent variable and the concept classes as independent variable, followed by pairwise contrasts between the EMMs for the concept classes. This will allow us to asses differences in sentiment intensity between the concept classes. We cannot perform planned contrasts by target polarity, because the descriptive concepts only have a neutral polarity, which leads to empty interaction levels and contrasts.

NEW: Even though polarity contrasts are not possible with descriptive concepts, we have to control that the remaining classes are distinct for negative and

positive concepts respectively. Hence, we should run a model without descriptive concepts, in order to measure polarity differences. H_5 expects that *positive* ethical, epistemic and legal thick concepts have significantly different sentiment averages, and the *negative* ones do so as as well among each other. **Note:** The model contains untransformed sentiment values as DP, and an interaction of polarity and concept class as IV.