Translation is a special case of language production resulting in a so-called “third code” that shows traces of the source language (SL) while trying to conform to the conventions of the target language (TL) (House 2010). There is a wealth of descriptive, corpus-based work on SL shining-through and TL normalization in written translation, converging on observations of skewed frequency distributions of selected linguistic features in translations compared to original (non-translated) productions (“translationese”). Existing computational approaches have applied automatic classification and clustering (Rabinovich/Wintner 2015, Volansky et al. 2015, Rubino et al. 2016) to tell translations from non-translations or to guess the SL of a given TL text. With few exceptions (He et al. 2016), computational accounts typically focus on written translation and start from handcrafted, often shallow features (e.g. sentence length, type-token ratio).

We here advocate an empirical approach using computational language models (LM) to detect and interpret linguistic effects of SL shining-through and TL normalization at linguistically more informative levels (lexis, grammar). Also, we are primarily interested in simultaneous interpreting because it poses fairly severe constraints on processing and working memory (Hyönä et al. 1995) and will thus show stronger shining through effects and weaker normalization effects than written translation and possibly other effects directly stemming from cognitive limitations.

Using a richly annotated version of the EuroParl Corpus (Karakanta et al. 2018) and a comparable corpus of interpreting transcripts, we compare professional, written translations with professional, simultaneous interpretations of European Parliament speeches from English into German. We build language models for English originals, German translations and interpretations and German originals and compare models in terms of perplexity/relative entropy as follows:

(a) to assess the relative, overall difference between translation and interpreting, we compare word-based and part-of-speech based ngram LMs for translations vs. interpretations in the target language;
(b) to assess effects of TL normalization, we compare word-based and part-of-speech based LMs of (i) translations vs. originals and (ii) interpretations vs. originals in the target language;
(c) to assess effects of SL shining-through, we compare part-of-speech based LMs of (i) translations vs. originals and (ii) interpretations vs. originals in the source language.

We find, for instance, that overall, interpretations exhibit less natural (more surprising) word order choices (due to SL shining-through) and tend to prefer paratactic structures, while translations tend towards hypotactic structures (a trace of TL normalization), indexed by unusual ngrams and the choice of particular conjunctions.