Despite the growing interest in the use of corpus material and methodologies in translation research, there is a lack of a systematic description of the various statistical tests that may be adapted from corpus linguistics in general for the purpose of translation research. This seminar will present results from a forthcoming edited volume designed to address this issue.

Firstly, Lidun Hareide and Knut Hofland describe through practical advice the compilation process of The Norwegian Spanish Parallel Corpus (NSPC) created at the University of Bergen (Norway), as well as preliminary findings from ongoing and planned research based on it. The corpus is primarily constructed for research in Translation Studies, and is built to be roughly comparable to the Spanish-English P-ACTRES corpus.

Next, Michael P. Oakes describes various ways of describing a text, whether in original or in translation, considering measures of vocabulary richness, how to find which linguistic features most typify a text, growth of vocabulary with text length, and whether or not the corpus is uniform in composition.

Finally, Meng Ji and Michael P. Oakes compare two translations of Cao Xueqin’s classic Chinese novel “Dream of the Red Chamber” using the whole range of commonly-used “bivariate” statistical texts for comparing one group (or corpus) with another.