Dialect comparison in the Anselm Corpus

Stefanie Dipper (Bochum)

In my talk, I present work on the Anselm Corpus. This corpus consists of more than 50 texts in Early New High German which are dialectal (diatopic) variants of the same underlying text, "Questions by Saint Anselm about the Lord's Passion". The corpus has been annotated semi-automatically at different levels: normalized and modernized wordforms, character-based alignments, POS, morphology, lemma. A subset of texts has been manually aligned for translation equivalents and annotated for keywords. I will show how the texts and their annotations can be used for dialect comparison at different linguistic levels.