

### Generating the *Stemma codicum* of Petrus Alfonsi's *Dialogus*

In this talk we present the manuscript tradition of Petrus Alfonsi's *Dialogus contra Iudaeos*, written around AD 1110. The author, a Spanish Jew who converted to Christianity, soon became a popular teacher due to his good knowledge of Arabic science and his acquaintance with contemporary Judaism and Islam. The manuscript tradition suggests that he wrote his *Dialogus* in northern France or possibly in southern England. The text was widely disseminated in the Middle Ages and there are over sixty more or less complete manuscripts known. In his dialogue the author discusses with his old Jewish alter-ego which of the three monotheistic religions is best.

Because the text was so widely disseminated within a century of its composition, constructing a stemma is difficult. Our oldest surviving manuscript dates to less than 20 years after the original composition of the text and yet it is already part of a clear-cut group among about half a dozen others. So we have here a textual tradition that bifurcated rapidly in the first half century of its existence. Apparently the text was spread mainly by the Benedictine and the Cistercian orders all over Western Europe. As usually in such cases of popular texts we have the problem of contamination among some manuscript(-branches).

Our approach to finding a stemma is a combination of philological and computational methods. The computation is based on an excerpt of some 500 words from the text of each manuscript from three different portions of the text. We introduce the notion of a metric, defining distances between samples from any two different manuscripts. These distances are to be calculated for each pair of samples and arranged in a matrix. Based on this distance matrix we then calculate an unrooted tree diagram using standard methods from molecular biology.

The description of a "distance" function between two manuscript texts should ideally include all our knowledge on scribal tradition, distinguishing trivial from non-trivial mistakes, with irreversible alteration having a much higher weight than reversible ones and trivial spelling variants no weight at all. The figures thus calculated will be an expression of the sum of the number of alterations that happened between the two samples in question. The distance is, of course, fixed at zero for word-by-word identical texts. Such an algorithmically implementable distance will mimic the work of a good traditional human editor as closely as possible. We have developed a simple algorithm capable of estimating whether a difference between texts is trivial or a non-trivial *leitfehler* (*error significativus*). Based on this, we can generate a reasonably realistic stemma. In a second step, we fine-tune the distance function manually by screening the *leitfehler* candidates suggested by the algorithm, resulting in a further improvement of the generated stemma.

For the tradition of the *Dialogus* we were thus able to determine one subgroup of the tradition with clearly secondary readings. Unfortunately the only modern editor of the text used a manuscript from this subgroup as his main witness. Apart from this subgroup we could differentiate eight groups and three to five manuscripts that do not fit well in any of the groups. The major problems in the stemma occur close to the archetype where traditional methods of textual criticism revealed contamination in at least one subgroup. Our results could be confirmed by finding *leitfehler* that separate our groups using traditional methods; our stemma will pave the way for an overdue critical edition of this text.