

Studia Stemmatologica IX

New Insights in Stemmatology, Bergen, 30th June - 1st July 2022



Collection of abstracts

Speakers

Teemu Roos, University of Helsinki

Armin Hoenen, Goethe University Frankfurt

Tiago Tresoldi, Uppsala University

Luigi Bambaci, University of Bologna

Tara Andrews, Tatevik Atayan and Anahit Safaryan, University of Vienna

Marko Halonen, The Ella and Georg Ehrnrooth Foundation

Elisa Cugliana, Ca' Foscari University of Venice

Bjarni Gunnar Ásgeirsson, University of Iceland

Caroline Macé, Univ. of Hamburg / Göttingen Acad. of Sciences and the Humanities

Philipp Roelli, University of Zurich

Elisabet Göransson, Lund University

Peter Robinson, University of Saskatchewan

Michael Stolz, University of Bern

Jamie Tehrani and Gessica Martini, Durham University

Tuomas Heikkilä, University of Helsinki

Jost Gippert, University of Hamburg / Centre for the Study of Manuscript Cultures

Organisers

Marina Buzzoni, Ca' Foscari University of Venice

Aidan Conti, University of Bergen

Odd Einar Haugen, University of Bergen

Teemu Roos

University of Helsinki / Department of Computer Science

The Computer Says “Maybe”

Embracing Uncertainty in Computer-Assisted Textual Scholarship

Computer-assisted methods have been gaining ground in stemmatology and in textual scholarship more generally. They can be helpful in transcribing, collating, and analysing textual traditions as well as presenting the results in the form of charts and diagrams. However, due to the lack of transparency of the methods, the scholar may sometimes have little choice but to blindly trust the outcome of the analysis without being able to realistically assess how uncertain it is – this amounts to treating the method as a “black box” where the data goes in and the results come out with no control over or knowledge of what happens inbetween.

One potential cure to the problem is to provide quantitative indicators of the uncertainty associated with the result. We refer to such indicators as uncertainty quantification (UQ). UQ can serve as a warning signal that helps the scholar avoid trusting the outcome of computer-assisted methods too much when such trust would be misplaced. Common UQ techniques include statistical confidence values such as p-values, Bayesian posterior distributions, confidence intervals, and bootstrap values.

To get an idea about how common the use of UQ is in digital scholarship, we carried out a small survey of the literature published. We collected all papers published in Digital Scholarship in the Humanities between 1/2010–11/2021 and that include the phrase “computer-assisted” in the abstract (n=20 papers). Out of the surveyed papers, 13 proposed and/or applied methods for which UQ can be considered relevant, and among these, five (5) provided some kind of uncertainty quantifications. That is, a majority of papers failed to provide UQ.

We provide some concrete suggestions for easy to use techniques for UQ especially in the context of stemmatology, and encourage the community to doubt anything the computer says – especially if it doesn't include the word “maybe”.

Armin Hoenen

Goethe University Frankfurt

DeepLearning in Stemmatalogy

While in many other domains neural network driven approaches have outperformed traditional or other machine learning approaches, deep learning has not yet been applied to stemmatology as far as the author is aware. With the outstanding improvements in almost all computational fields and tasks not only has deep learning demonstrated to be an approach that should not be ignored, at the same time feature engineering of classical machine learning and annotation has been rendered ever more superfluous (compare e.g. the ability of BERT models to solve anaphora resolution solely on the basis of unannotated input data, Joshi et al. 2019). One drawback is however that this technology is data hungry and computationally intense. These issues have and will continue to complicate an application in stemmatology where to date no comprehensive international electronic database for published stemmata with their underlying collations has been created.

Two ways to mitigate this are the use of related data such as multimedia phylogenies as in Marmerola et al. (2016), who are among the few scholars to have applied classical machine learning in this branch. The second and among philologists probably and for understandable reasons infamous approach would be the use of simulated data. Within this second approach, either manually simulated data (as in the artificial datasets, Roos & Heikkilä 2009, Spencer et al. 2004, Baret et al. 2006, Hoenen 2015a) or computationally simulated data (to date only used in some studies on theoretical stemmatology, compare Weitzmann 1982, Flight 1994, Hoenen 2016) can be produced while only the automatic approach would be able to generate as much data as allegedly needed for performant deep learning.

In this contribution, on the basis of previous research a large artificial simulated dataset combining collations with true stemmata as a basis for subsequent applications of deep learning will be created and made freely available. Furthermore, a deep learning approach to stemmatology will be executed similar to bio-informatical precedents such as Suvorov et al. (2019), but domain adapted. Finally, the problem of large variable tree spaces and a mitigation by using an approach of reinforcement learning much as in AlphaGo Zero (see e.g. Holcomb et al. 2018) will be outlined. Maybe if understanding stemmatology as “the philologists game”, deep learning approaches are thinkable which could be applied to any stemmatological task once trained.

References

- Baret, P., Macé, C., and P. Robinson (2006). Testing methods on an artificially created textual tradition. In C. Macé, P. V. Baret, A. Bozzi, & L. Cignoni (eds.) *The Evolution of Texts: Confronting Stemmatalogical and Genetical Methods. Proceedings of the International Workshop held in Louvain-la-Neuve on September 1-2, 2004* (Linguistica

- Computazionale, 24-25). Pisa-Roma. Istituti Editoriali e Poligrafici Internazionali: 255-281.
- Flight, C. (1994). A complete theoretical framework for stemmatic analysis. *Manuscripta*, 38(2): 95-115.
- Hoenen, A. (2014). Simulation of Scribal Letter Substitution. In T. Andrews & C. Macé (eds.) *Analysis of Ancient and Medieval Texts and Manuscripts: Digital Approaches* (Lectio. Studies in the transmission of Texts and Ideas, 1). Turnhout: Brepols: 119-139.
- Hoenen, A. (2015a). Das artifizielle Manuskriptkorpus TASCHE. In DHd 2015 - *Von Daten zu Erkenntnissen - Book of abstracts*. Dhhd.
- Hoenen, A. (2015b). Simulating Misreading. In C. Biemann et al. (eds.) *Natural Language Processing and Information Systems*. Cham: Springer, 385-389.
- Hoenen, A. (2016). Silva Portentosissima Computer-Assisted Reflections on Bifurcativity in Stemmas. In *Digital Humanities 2016: Conference Abstracts*. Jagiellonian University & Pedagogical University: 557-560. <http://dh2016.adho.org/abstracts/311>.
- Hoenen, A. (2018). Multi Modal Distance - An Approach to Stemma Generation with Weighting. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.
- Holcomb, S. D., Porter, W. K., Ault, S. V., Mao, G., & Wang, J. (2018). Overview on deepmind and its alphago zero ai. In *Proceedings of the 2018 international conference on big data and education*: 67-71. <https://doi.org/10.1145/3206157.3206174>.
- Joshi, M, Levy, O., Zettlemoyer, L. & Weld, D. (2019). BERT for Coreference Resolution: Baselines and Analysis. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*: 5803-5808.
- Marmerola, G. D., Oikawa, M. A., Dias, Z., Goldenstein, S., & Rocha, A. (2016). On the reconstruction of text phylogeny trees: evaluation and analysis of textual relationships. *PLoS one*, 11(12): e0167822.
- Roos, T. & Heikkilä, T. (2009). Evaluating methods for computer-assisted stemmatology using artificial bench-mark data sets. *Literary and Linguistic Computing*, 24: 417-433.
- Spencer, M., Davidson, E. A., Barbrook, A. C., and Howe, C. J. (2004). Phylogenetics of artificial manuscripts. *Journal of Theoretical Biology*, 227: 503-511.
- Suvorov, A., Hochuli, J., & Schrider, D. R. (2020). Accurate inference of tree topologies from multiple sequence alignments using deep learning. *Systematic Biology*, 69(2): 221-233.
- Weitzman, M. P. (1982). Computer simulation of the development of manuscript traditions. *ALLC Bulletin. Association for Library and Linguistic Computing* Bangor 10(2): 55-59.

Tiago Tresoldi

Uppsala University / Department of Linguistics and Philology

Illustrating Bayesian inference for stemmatology by example

The use of computational methods and tools in stemmatology has increasingly called the attention of the community, as illustrated by the *Handbook of Stemmatalogy* (Roelli, 2020) devoting a full chapter to them (van Zundert et al., 2020). Among these methods we find the Bayesian inference with Markov Chain Monte Carlo (MCMC) sampling, provided by tools such as MrBayes (Ronquist et al., 2012), RevBayes (Höhna et al., 2016), and BEAST2 (Bouckaert et al., 2014; Maurits et al., 2017), of great prominence among evolutionary studies in genetics and historical linguistics (Greenhill et al., 2020) but not as common in the study of textual evolution as tools like Phylip (Felsenstein, 1989), PAUP* (Swofford, 1998), and SplitsTree (Huson, 1998).

This workshop will extend the exposition in the above chapter by demonstrating the steps and tools involved in a stemmatological Bayesian inference analysis with BEAST2 (such as from Rambaut, 2010, and Bouckaert, 2010). It is planned so that philologists can familiarise with an “analysis pipeline”, highlighting decisions that a phylogeneticist needs to make when building a “model” and how they can affect the results. We will analyse the same dataset in a handful of different configurations, highlighting the limits and advantages of the method for each type of tradition and textual transmission in general. Audience participation throughout the presentation will be welcome. All data, scripts, the BEAST2 XML models, and results will be made available before the workshop. Nonetheless, the only requirement for the audience will be to have read van Zundert et al. (2020); it will not be necessary or assumed that the audience will reproduce or run any analysis during the presentation.

The simple and brief artificial tradition by Guillaumin (2020), presented while discussing the limits of digital methods, will be used during the whole workshop. It is a good “synthetic” tradition not only because the public will be familiar with it but also because of its size, the attention to the main difficulties of digital stomatology, and an available “correct” stemma (p. 342) with which the individual results can be compared.

The topics explored in the workshop will be:

- Common workflow of phylogenetic inference, analogies between textual and biological evolution
- Demonstration of Bayesian MCMC phylogenetic inferences
 - Construction of phylogenetic “characters” by annotation or automatic clustering (e.g., Dekker, 2011)
 - Calibrations and monophyletic constraints
 - Tree priors (“uniform”, “yule”, “coalescent”, and “birth-death”)
 - Evolutionary models (“covarion”, “BSVS” and “Lewis-Mk”)
 - Evolutionary clock types (“strict” and “relaxed”)
- Discussion on limits of current methods for stemmatology, alternatives, phylogenetic networks

References

- Bouckaert, R. 2010. "DensiTree: making sense of sets of phylogenetic trees". *Bioinformatics* 26(10): 1372–1373. <https://doi.org/10.1093/bioinformatics/btq110>.
- Bouckaert, R.; Heled, J.; Kühnert, D.; Vaughan, T.; Wu, C. H.; Xie, D. et al. 2014. "BEAST 2: A Software Platform for Bayesian Evolutionary Analysis". *PLoS Comput Biol.* 10(4). <https://doi.org/10.1371/journal.pcbi.1003537>.
- Dekker, R. H. and Middell, G., 2011. Computer-Supported Collation with CollateX: Managing Textual Variance in an Environment with Varying Requirements. *Supporting Digital Humanities 2011*. University of Copenhagen, Denmark. 17-18 November 2011.
- Felsenstein, J. 1989. "PHYLIP Phylogeny Inference Package (Version 3.2)". *Cladistics* 5 (2): 164–166. <https://doi.org/10.1111/j.1096-0031.1989.tb00562.x>.
- Greenhill, S. J.; Heggarty, P. and Gray, R. D. 2020. "Bayesian Phylolinguistics". In Janda, R. D.; Joseph, B. D. and Vance, B. S., *The Handbook of Historical Linguistics II*. Hoboken, NJ, USA: Wiley. <https://doi.org/10.1002/9781118732168.ch11>
- Guillaumin, J.-B. 2020. "Criticism of digital methods". In Roelli, P. (ed.). *Handbook of Stemmataology*, Berlin, Boston: De Gruyter.
- Höhna, Landis, Heath, Boussau, Lartillot, Moore, Huelsenbeck, Ronquist. 2016. RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Syst. Biol.*, 65: 726-736.
- Huson, D. H. 1998. "SplitsTree: Analyzing and Visualizing Evolutionary Data". *Bioinformatics* 14 (1): 68–73. <https://doi.org/10.1093/bioinformatics/14.1.68>.
- Maurits, L.; Forkel, R.; Kaiping, G. A. and Atkinson, Q. D. 2017. "BEASTling: A software tool for linguistic phylogenetics using BEAST 2". *PLoS ONE* 12(8): e0180908. <https://doi.org/10.1371/journal.pone.0180908>.
- Rambaut, A. 2010. *FigTree v1.3.1*. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh.
- Roelli, P. (ed.). 2020. *Handbook of Stemmataology*, Berlin, Boston: De Gruyter.
- Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D. L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M. A. And Huelsenbeck, J.P. 2012. "MRBAYES 3.2: Efficient Bayesian phylogenetic inference and model selection across a large model space". *Syst. Biol.* 61:539-542. <https://doi.org/10.1093/sysbio/sys029>.
- Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014 May 1;30(9):1312-3 <http://doi.org/10.1093/bioinformatics/btu033>.
- Swofford, D. L. 1998. PAUP* 4.0: Phylogenetic Analysis Using Parsimony (* and OtherMethods). Version 4.0. CD-ROM. Newer versions available at <http://paup.phylosolutions.com>.
- van Zundert, J.; Hoenen, A.; Manafzadeh, S.; Staedler, Y. M.; Roos, T.; Guillaumin, J.-B. 2020. "Computational methods and tools". In Roelli, P. (ed.). *Handbook of Stemmataology*, Berlin, Boston: De Gruyter.

Luigi Bambaci

University of Bologna

Applying cladistics to authoritative texts: The case of the Hebrew Bible

The medieval tradition of the Hebrew Old Testament or Hebrew Bible (hb) poses a series of important problems to philologists dealing with stemmatology: first, the contamination typical of open traditions; second, a lack of significant common errors, which renders Lachmann's method impracticable; and finally, a process of 'controlled transmission' [15, 14], which led scribes to suppress readings deviating from the *textus receptus* (tr).

These three features – contamination, lack of common errors, and hegemony of the tr – have, on one side, given rise to the impression of a substantial 'fixity' of the hb text [13] and, on the other, fed the conviction that it is impossible to organize Hebrew manuscripts into textual families by means of traditional philological methods [2, 5].

In contrast to that of the New Testament, which is similar in certain respects [1], the hbtradition has been the subject of few attempts at stemmatic classification. The most recent [10, 3] base themselves on clustering algorithms, while none – to the best of our knowledge – take into account the genealogical model.

We propose to describe an experiment in stemmatic analysis using cladistics, one of the most important methods in computer-assisted stemmatology [8], and taking as our case study the tradition of the book of Qohelet, also known as Ecclesiastes. In particular, we took the late 18th-century edition of Benjamin Kennicott [6] and examined the variants from textual witnesses, for the most part medieval codices, collated by him.

The method we used consisted of four steps: (1) construction of a data matrix from a file encoded in xml-tei; (2) application of the Maximum Parsimony criterion [7] as implemented in paup (<https://paup.phylosolutions.com>) (3) ancestral reconstruction, for identifying so-called characteristic variants [9]; (4) qualitative evaluation of the groupings, to establish their validity.

As a result, we were in fact able to distinguish groups of witnesses, divided according to the distribution of ancestral variants. Some of these seem likely, on the basis of both external criteria and the number or type of variants in common. Others, however, are problematic, since either the quantity or the quality of shared variants does not allow us to ascertain a genealogical kinship. The paucity of genealogically informative variants and in particular of characteristic variants seems to derive, in effect, from the absence of clear patterns of agreement, a phenomenon which is also typical of other textual traditions [4] and which represents, together with contamination and coincident variation, a challenge for (computational) stemmatology and for traditional methods of validating the reliability of *stemmata codicum* [12].

In light of the results obtained, we intend here to confront the implications of these problematic aspects of quantitative analysis, seeking to evaluate to what extent they affect the possibility of defining *stemmata* of the medieval tradition of the hb.

References

- [1] J. R. Adair. Old and New in Textual Criticism: Similarities, Differences, and Prospects for Cooperation. *TC: A Journal of Biblical Textual Criticism*, 1996.
- [2] D. Barthélemy. Les manuscrits médiévaux et le texte tибérien classique. In *Critique textuelle de l'Ancien Testament, 3. Ézéchiel, Daniel et les 12 Prophètes*, volume 3 of *Orbis Biblicus et Orientalis*, pages xix–xcvi. Éditions Universitaires/Vandenhoeck & Ruprecht, Fri bourg/Göttingen, 1992.
- [3] P. G. Borbone. *Il Libro del profeta Osea. Edizione critica del testo ebraico*. Zamorani, Torino, 1990.
- [4] B. Bordalejo and P. Robinson. Manuscripts with Few Shared Significant Variants. *Ecdotica*, 15:37–65, 2018.
- [5] B. Chiesa. *Filologia storica della Bibbia ebraica*, volume II of *Studi Biblici*. Paideia, Brescia, 2000.
- [6] B. Kennicott. *Vetus Testamentum Hebraicum cum variis lectionibus*, volume 2. Oxford, 1778.
- [7] A.-C. Lantin, P. Baret, and C. Macé. Phylogenetic Analysis of Gregory of Nazianzus' Homily 27. In *Le poids des mots. Actes des 7èmes Journées Internationales d'Analyse statistique des Données Textuelles*, volume 2, pages 700–7, Louvain-la-Neuve, 2004.
- [8] P. Robinson. Computer-Assisted Stemmatic Analysis and 'Best-Text' Historical Editing. In P. T. v. Reenen, M. v. Mulken, and J. Dyk, editors, *Studies in Stemmatology*, volume I, pages 71–104. John Benjamins Publishing, 1996.
- [9] P. Robinson. A Stemmatic Analysis of the Fifteenth-Century Witnesses to The Wife of Bath's Prologue. In P. Robinson and N. F. Blake, editors, *Canterbury Tales Project Occasional Papers*, volume II, pages 69–132, London, 1997.
- [10] P. Sacchi. Analisi quantitativa della tradizione medievale del testo ebraico della Bibbia secondo le collazioni del De Rossi. *Oriens Antiquus*, 12:1–13, 1973.
- [11] B. J. P. Salemans. *Building Stemmas with the Computer in a Cladistic, Neo-Lachmannian, Way: The Case of Fourteen Text Versions of Lanseloet Van Denemerken*. University Press, 2000.
- [12] M. Spencer, B. Bordalejo, P. Robinson, and C. Howe. *How Reliable is a Stemma? An Analysis of Chaucer's Miller's Tale*, volume 18. Nov. 2003.
- [13] S. Talmon. The Old Testament Text. In C. F. Evans and P. R. Ackroyd, editors, *The Cambridge History of the Bible: Volume 1: From the Beginnings to Jerome*, volume 1 of *The Cambridge History of the Bible*, pages 159–199. Cambridge University Press, Cambridge, 1970.
- [14] R. D. Weis. Lower Criticism: Studies in the Masoretic Text and the Ancient Versions of the Old Testament as Means of Textual Criticism. In M. Saebo, editor, *Hebrew Bible / Old Testament The History of Its Interpretation*, volume 3, pages 346–92. Vandenhoeck & Ruprecht, Göttingen, 1996.
- [15] J. W. Wevers. A Study in the Hebrew Variants in the Books of Kings. *Zeitschrift für die alttestamentliche Wissenschaft*, 61(1):43–76, 1948.

Tara Andrews, Tatevik Atayan and Anahit Safaryan

University of Vienna, Independent Scholar, and University of Vienna

Combining classical and computational approaches to the construction of a stemma for the *Chronicle of Matthew of Edessa*

The *Chronicle* of Matthew of Edessa, an Armenian historical work written in the first half of the twelfth century, provides an outstanding example of a textual tradition that does not easily lend itself to any one form of stemmatic analysis (Andrews 2016, 164–71). The overall work is relatively long, at some 80,000 words, and is transmitted in at least 35 manuscripts. The text they carry is relatively stable; although some of the manuscripts carry truncated texts, the magnitude of variation is reasonably small. The original text was written by an Armenian resident of the city of Edessa, in the early stages of the development of the late-medieval Cilician dialect, of which the only in-depth study is that of Karst (1901); although the text is written in *grabar* (classical Armenian), it shows the influence of its author's idiom and thus makes grammatical correctness a dangerous ground on which to analyze priority of readings. Complicating the situation even further is the fact that all extant copies of the text date from early modern times and were written in a variety of locations across the Armenian diaspora, within communities of clerics who maintained substantial contact with each other.

In this paper we present an updated and expanded version of the stemma for the *Chronicle*, an initial version of which was published a decade ago (Andrews 2009). For the new analysis, which takes into account eleven witnesses that were previously unavailable to us, we carried out independent “classical” and “computational” analyses and then compared the results. The classical analysis relies on paratextual features of the witness texts such as colophons, outer text structure, and substantial gaps as well as on close analysis of selected variant text passages. The computational analysis draws on three well-known tree-generation methods (Pars: Felsenstein 2013; RHM: Roos, Heikkilä, and Myllymäki 2006; NeighborNet: Bryant and Moulton 2004) to construct a separate initial hypothesis for the genealogical relationship of the manuscripts to each other. These methods have then been compared and reconciled with each other, resulting in a substantial re-thinking and re-orientation of the initial stemma, with tangible consequences for the future re-construction of the text.

Ultimately, we present our work not only for the sake of this specific text, but also in service to the idea that philologists can and should make use of all the available tools, both classical and computational, and that they need not be at odds with each other as is occasionally implied in scholarly polemic within the field.

References

Andrews, Tara L. 2009. “Prolegomena to a Critical Edition of the Chronicle of Matthew of Edessa, with a Discussion of Computer-Aided Methods Used to Edit the Text.” D.Phil. thesis, Oxford: University of Oxford.

- Andrews, Tara L. 2016. *Matt'ēos Urhayec'i and His Chronicle: History as Apocalypse in a Crossroads of Cultures*. Medieval Mediterranean 108. Leiden; Boston: Brill.
- Bryant, David, and Vincent Moulton. 2004. "Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks." *Molecular Biology and Evolution* 21 (2): 255–65.
- Felsenstein, Joseph. 2013. *PHYLIP (Phylogeny Inference Package) (version 3.695)*. Seattle: Department of Genome Sciences, University of Washington.
<https://evolution.genetics.washington.edu/phylip.html>.
- Karst, Josef. 1901. *Historische Grammatik Des Kilikisch-Armenischen*. Strassburg: Verlag von Karl J. Trubner.
- Roos, Teemu, Tuomas Heikkilä, and Petri Myllymäki. 2006. "A Compression-Based Method for Stemmatic Analysis." In *Frontiers in Artificial Intelligence and Applications*, 141: 805–6.

Marko Halonen

The Ella and Georg Ehrnrooth Foundation

Using digital stemmatological methods in creating a critical edition of a 16th-century chronicle

Bishop Paulus Juusten's (c. 1520–1575) chronicle, *Catalogus et ordinaria successio episcoporum Finlandensium*, is one of the most important historical sources concerning the Late Middle Ages in Sweden, particularly of its eastern parts, which today are mainly part of Finland. The manuscript tradition has 15 witnesses, the latest (Ms R) was discovered by the author in 2012. The chronicle has been edited/translated four times before: C. von Nettelblatt (1728)¹, H.G. Porthan (1799)², W. Schmidt (1943)³, and S. Heininen (1988)⁴. After having studied the chronicle earlier,⁵ I am now working towards a new critical edition. This is necessary for several reasons: The newly discovered manuscript changes the value of the witnesses, digital stemmatology was not available earlier, and the possibilities of publishing have changed significantly.

The editions by Porthan and Schmidt were based on what could be called the Renaissance-method of textual criticism. This means trying to fill in the gaps or 'errors' found in manuscripts by replacing them with the variant reading that seems the best. This relies on the editor's expertise concerning the language and historical context. Instead, Simo Heininen's edition is vaguely based on the Lachmannian method, but he failed to be consistent, which resulted in the 1988 edition becoming a hybrid between the selective Renaissance, and reconstructive Lachmannian methods. The first editor, C. von Nettelblatt applied the best manuscript - also as known as Bédierist approach. By selecting one manuscript he could sure that something authentic was shown to the reader, even if this is probably not exactly what Juusten had written.

Digitalization offers a multitude of possibilities for academic editing, which should be fully taken advantage of. Therefore, I am planning to publish all manuscripts and their transcriptions online in order to make the editorial process absolutely transparent to the

¹ Christian von Nettelblatt. *Schwedische Bibliothec*. Leipzig 1728.

² *M. Pauli Juusten Chronicon episcoporum Finlandensium*. [Edited by H. G. Porthan]. Officina Frenckelliana, Turku 1799.

³ *Suomen Piispain Kronikka*. [Edited by Wolfgang Schmidt, translation by Helmer Winter]. Suomentaja, Rauma 1943.

⁴ *Paulus Juusten Catalogus et Ordinaria Successio episcoporum Finlandensium*. [Edited by Simo Heininen.] Suomen kirkkohistoriallisen seuran toimituksia, Helsinki 1988.

⁵ Marko Halonen (2012). *Stemmatology of a 16th Century Chronicle: a comparison of traditional and computer-assisted methods for creating stemmata of Paulus Juusten's Catalogus et Ordinaria Successio Episcoporum Finlandensium*. Master's thesis, University of Helsinki, <https://helda.helsinki.fi/handle/10138/38059>, Computer-assisted stemmatology in studying Paulus Juusten's 16th-century chronicle *Catalogus et ordinaria successio Episcoporum Finlandensium*. *Digital Scholarship in the Humanities*. Volume 31, Issue 3, September 2016, pp. 578–593. Oxford University Press. <https://doi.org/10.1093/llc/fqv004>

reader. There are several platforms and programs which allow comparing two or more manuscripts and transcriptions simultaneously.⁶ I am also planning to publish a translation, and possibly several translations (English, Finnish, Swedish) as an e-book, and if possibly, as an audiobook.

However, after being very much involved with digital humanities for the past decade both in terms of academic research, high school teaching and digital editing, I must admit that the rapid development of methods, platforms and applications has also demonstrated the many pitfalls that often accompany new technology. These range from screen time to the loss of quality in favor of quantity. Ever-changing technical requirements, stylistic tastes and methods make it absolutely necessary to publish the new critical edition as a book as well: *Littera scripta manet*.

The limits of a book as a format can actually be turned into advantages. The impossibility of being able to show all manuscripts will actually force me to choose one as a base text. This decision must be well argued in the introduction. The basis of this argument must be a digital stemmatological analysis based on the entire tradition with all the witnesses. However, after studying a super-contaminated tradition of the medieval calendar, I think one should abandon the idea of trying to discover the original and focus instead on describing the variants and their history.⁷

The manuscripts must be transcribed in a systematic fashion. The challenge of many critical editions is that the transcription is done anachronistically (and without showing the original manuscripts), by using letters and punctuation which did not exist at the time.⁸ The transcription should therefore be as close to late 16th-century style as possible, although this is not as much a problem with Latin as with Greek. Thirdly, the edition should include a translation, the purpose of which is to demonstrate how the message of the original manuscripts is perceived by the translator. Fourthly, the edition must contain an extensive critical commentary, which will be possible due to the aimed folio-size page, and also using innovative (or actually re-discovered) ways of presenting the information, such as margins and footnotes.⁹

⁶ E.g. the Textual Communities project by the University of Saskatchewan.
<https://textualcommunities.org/app>

⁷ Marko Halonen (2020). Aspects of medieval Nordic calendars : a qualitative, quantitative and phylomemetic study. Doctoral dissertation, University of Helsinki
<http://urn.fi/URN:ISBN:978-951-51-5975-5>.

⁸ For example, probably the most studied book in the world, the New Testament, has been for decades presented with a form of Greek which certainly was not used in the first centuries A-D. *Novum Testamentum Graece*. Ed. Nestle & Aland XXVII, 1995.

⁹ One way of doing this is e.g. in the edition of the Hebrew old testament. *Biblia Hebraica Stuttgartensia*, 1997.

Elisa Cugliana

University of Cologne / Cologne Center for eHumanities

***Realia* and other clues for stemmatological wayfinding**

The case of the Early New High German Marco Polo and its Tuscan model

Although not everybody took it seriously, the travel narrative by Marco Polo and Rustichello da Pisa enjoyed immense success in the Middle Ages. The consequence was a proliferation of copies and copies of copies of the text. Indeed, more than 140 manuscripts, produced within a time span of two centuries, have survived until the present day. To make matters worse (or, to put it another way, more interesting), the text was translated into more than a dozen languages and often profoundly modified, connotating its transmission with an extreme *mouvance*. Consequently, there are still numerous enigmas to be solved, especially as far as the *stemma codicum* of Marco and Rustichello's work is concerned.

As a matter of fact, some of its minor branches run the risk of being overlooked by the scholarly community, busy solving problems at the top of the *stemma*. For instance, the German translation DI of the work has not enjoyed much scholarly attention and it is now being edited for the first time in the context of a PhD project. The focus of the paper is on the methodology used to address the hazy constellation of DI and its Tuscan model TB: in this respect, the workflow included both quantitative methods (especially for the collation, performed with *CollateX*) and qualitative ones. Specifically, it will be shown that named entities and *realia* play an important role in the collation process, serving as “index fossils” (Reginato, 2016) in the quest for *Leitfehler*. Some linguistic features of these words make them particularly suitable for stemmatic analyses, as they are often prone to fulfil the criteria of uniqueness and irreversibility of the error required in stemmatology. This is especially due to their exceptional semantics and, as a result of it, their resistance to translation (and transcription).

Indeed, by applying this methodology it was possible to call into question the position of DI within the TB configuration envisioned by Benedetto (1928) and formulate a new proposal. In this novel hypothesis, which is of interest not only for Marco Polo studies, but also for research in stemmatology, DI is collocated in a different branch of the TB *stemma*. Such a change is not of little value, in particular because the Tuscan witness that served as model for the German translation was lost, meaning that DI is extremely relevant for the establishment of a critical text for TB, whose original is, alas, also lost.

References

Benedetto, Luigi F., Marco Polo, *Il “Milione”*, prima edizione integrale a cura di L.F. Benedetto, Firenze 1928.

Cruse, Mark, “A quantitative analysis of toponyms in a manuscript of Marco Polo's *Devisement du Monde*”, *Speculum* 92, S1 (2017), pp. 247-264.

- Dekker, Ronald H. and Gregor Middell, *CollateX – Software for Collating Textual Sources*, <<https://collatex.net>> [12/02/21
- Gadrat Ouerfelli, Christine, *Lire Marco Polo au Moyen Age. Traduction, diffusion et réception du Devisement du monde*, Turnhout 2015.
- Molino, Jean, “Le nom propre dans la langue”, *Langages* 16, 66 (1982), pp. 5-20.
- Reginato, Irene, “I nomi propri come ‘fossili-guida’ nello studio filologico-linguistico di un testo. Il caso della Versione K del Devisement du Monde” in Maria P. Arpioni, Arianna Ceschin and Gaia Tomazzoli (eds), *Nomina sunt...? L’onomastica tra ermeneutica, storia della lingua e comparatistica*, Venezia 2016, pp. 65-76.
- Simion, Samuela, “Tradizioni attive e ipertesti. Ramusio ‘editore’ del Milione”, *Quaderni veneti* 6, 1 (2017), pp. 9-30.

Bjarni Gunnar Ásgeirsson

University of Iceland

The top of the stemma – a case study

In 1928, Joseph Bédier claimed that out of 110 stemmata for Old French texts, 105 were bifid. As Bédier did not specify from where he got his numbers, Arrigo Castellani made his own survey and found that 82.5 per cent of Old French stemmata published before 1928 were bifid. More recently, Odd Einar Haugen (2016) has surveyed stemmata published with studies and editions of Old Norse texts in the series *Bibliotheca Arnamagnæana* and *Editiones Arnamagnæanæ*, published in Copenhagen between 1938 and 2013, and found that the numbers are almost identical to Castellani's: 83 per cent of the Old Norse stemmata are bifid; the rest are split in three or more branches.

The preponderance of bifid stemmata puzzled Bédier, who suggested that editors of Old French texts were reducing multi-branched stemmata until they were left with just two families in order to have more freedom to reconstruct the text as they saw fit, or, alternatively, that they inadvertently divided the witnesses into two families until nothing was left, seeing conjunctive errors where there were none. This phenomenon has been debated ever since, and arguments have been made that there are historical and mathematical reasons for why two-branched stemmata are more common (Roelli 2020). As the scholars studying the Old Norse texts were not attempting to reconstruct a lost archetype, Haugen found no reason to suspect that they had manipulated their stemmata. However, he did find it possible that they were affected by “the force of dichotomy” and that there could be a “tendency to divide the material till the end of the line, to see splits where there may be no splits, to look for divergence rather than for unity” (Haugen 2016: 608).

In another series of Old Norse text editions, *Íslenzk fornrit*, we do come across attempts to reconstruct lost archetypes. In his edition of *Brennu-Njáls saga*, Einar Ól. Sveinsson split the eighteen medieval witnesses (mostly fragments) into three groups: X, Y and Z. He found that there was “a special affinity between Y and Z” and so he assumed that they shared an intermediary manuscript (called V) and produced a bifid stemma. Einar's arguments for the existence of V are mostly based on fifteen readings (in a text of about 100,000 words), where he found that X had readings “better” than V; but as Bédier noted, the top of the stemma is always the hardest to pinpoint, and applying some subjectivity is unavoidable. In this paper, I will examine the stemmata of *Brennu-Njáls saga*, re-evaluate Einar's work, and argue that the evidence in favour of the hypothetical V may be too meagre to sustain the proposed bifid stemma, and that Einar may have been affected by “the force of dichotomy”.

References

Einar Ól. Sveinsson. *The Manuscript Tradition of Njáls saga*. Studia Islandica 13. Reykjavík: H.F. Leiftur, 1953.

Haugen, Odd Einar. "The *silva portentosa* of stemmatology: Bifurcation in the recension of Old Norse manuscripts." *Digital Scholarship in the Humanities*, 31, 3 (September 2016), 594–610. <https://doi.org/10.1093/llc/fqv002>

Roelli, Philipp, "Definition of stemma and archetype." *Handbook of Stemmatology: History, Methodology, Digital Approaches*, ed. by Philipp Roelli, 205–225. Berlin, Boston: De Gruyter, 2020. <https://doi.org/10.1515/9783110684384>

Caroline Macé

University of Hamburg / Göttingen Academy of Sciences and the Humanities

Reconstructing the top of a stemma using an outgroup and its consequences for the edition

The *stemma codicum* does not only have a practical mean as visualising tool, but also a symbolic value as warrant of scholarly standard. As symbol, it may remain a purely decorative accessory in a philological introduction that only few are going to read. In that case, it does not matter much for the edition whether the stemma is accurate or not (Menestò 1981). As a touchstone to assess the value of a stemma, the use of an “out-group” (especially translations) proves essential. Taking a few examples in (Greek) Patristic literature, I will show how important it is to use ancient translations and how much a change in the topography of the stemma, especially at the top, does affect the edited text.

(1) Physiologus (anonymous, 3rd cent.)

Sbordone 1936 provided the first classification of 70 Greek manuscripts (11th–17th cent.) into three recensions and stemmata for each recension. Sbordone could not use the oldest translations of the first recension: Armenian (5th or 6th cent.), Ethiopic (7th or 8th cent.), Latin (before the 8th cent.), Syriac (6th or 7th cent.). A comparison of the Greek text with the translations leads to the conclusion that Sbordone’s stemma, divided into four branches, is wrong (Macé and Gippert 2021). Sbordone’s assessment of manuscript “M” as the oldest and “best” manuscript is equally erroneous. Two editions of the *Physiologus* appeared after Sbordone’s. The first one (Offermans 1966) is a monotypic edition of a manuscript older than “M” but belonging to the same secondary family (branch 1). The second (Kaimakis 1974) is a synoptic publication of the texts of Sbordone’s branches 2-4, without any reconsideration of Sbordone’s stemma.

(2) Gregory of Nazianzus’ Homilies 10 and 12 (c. 380)

Mossay 2006 used phylogenetic methods and compared the c. 120 Greek manuscripts (9th–16th cent.) with Syriac (c. 625) and Georgian (end of 11th cent.) translations. However, the collations were often faulty, the methods were applied without really understanding them and the translations were often misinterpreted. As a result, the stemma has no value and the variants are chosen *ad libitum*, ending up in an arbitrary eclectic text.

(3) Pseudo-Dionysius Areopagita’s Letter on the death of the Apostle Paul (5th–7th cent.)

This letter exists in Arabic, Armenian, Ethiopic, Georgian, Latin, Syriac and Early Modern High German, reflecting two different Greek recensions, both lost (Macé et al. 2021). The second recension (probably created in a Greek monastery in Rome in the 7th cent.) is extant in 8 Georgian (10th–15th cent.) and 118 Latin (13th–15th cent.) manuscripts. Because I had misunderstood the nature of the relationship between the Georgian and the Latin texts, my stemma of the Georgian manuscripts was at first erroneous and contradictory.

References

- Kaimakis, Denis (1974), *Der Physiologus nach der ersten Redaktion* (Beiträge zur klassischen Philologie, 63), Meisenheim am Glan: Anton Hain.
- Macé Caroline and Jost Gippert (2021), *The Multilingual Physiologus. Studies in the Oldest Greek Recension and its Translations*, ed. by (Instrumenta Patristica et Mediaevalia, 84), Turnhout: Brepols, 2021.
- Macé, Caroline, Ekkehard Mühlenberg, Michael Muhtreich and Christine Wulf (2021), *Corpus Dionysiacum*, vol. III/1: *Epistola ad Timotheum de morte apostolorum Petri et Pauli* (Patristische Texte und Studien, 79), Berlin: De Gruyter.
- Menestò, Ernesto (1981), 'Problemi di stemmatica in edizioni critiche', in *La cultura in Italia fra tardo antico e alto medioevo*, Rome: Herder Ed. e Libreria, vol. 1, 443–466.
- Mossay, Justin (2006), *Sancti Gregorii Nazianzeni Opera. Versio Graeca*, I: *Orationes X et XII* (Corpus Christianorum. Series Graeca, 64 = Corpus Nazianzenum, 22), Turnhout: Brepols.
- Offermans, Dieter (1966), *Der Physiologus nach den Handschriften G und M* (Beiträge zur klassischen Philologie, 22), Meisenheim am Glan: Anton Hain.
- Sbordone, Francesco (1936), *Physiologus*, Rome: Societas 'Dante Alighieri – Albrighi, Segati et C.'.

Philipp Roelli

University of Zurich / Department of Greek and Latin Philology

Too complicated for digital tools to be of much help?

The *Liber Aurelii* and Pseudo-Ptolemy's *Centiloquium*

This talk compares two difficult cases of Latin textual transmissions and the question how far currently available digital tools are helpful in the reconstruction of the stemma and to provide clues how to best edit the texts.

The anonymous so-called *Liber Aurelii* is a late antique Latin medical work on acute illnesses that goes back to lost Greek sources, mostly from the Methodic school and Soranus of Ephesus. The work was quite popular, there are more than a dozen extant manuscripts, the oldest of which dating from the early ninth century. Several layers of textual “erosion” and attempts to re-establish it have partially survived. A popular eleventh century medical compendium (65 manuscripts) quotes almost the entire text, its author Gariopontus still had a more comprehensible manuscript available. Already before him, an anonymous abbreviator worked on the text, shortened it considerably and made its content much clearer. I edited the text in 2021 for the first time critically, all three recensions are edited in parallel.

The second text is the Latin version of Pseudo-Ptolemy's *Centiloquium*, a collection of aphorisms on astrology, currently being edited by my colleague Emanuele Rovati. This short text was translated from the Arabic, apparently by Plato of Tivoli (12th century). It is extant in more than 100 manuscripts in two different versions, one reworked by Gerard of Cremona. The lost Greek original was translated into Arabic, from whence it was translated by Plato. The text was printed in 1484. The talk will show that – despite the many witnesses – the situation is less complex and more amenable to computer-aided study. Rovati's forthcoming edition will probably print the archetypal text with an extra apparatus for Gerard's changes.

The use of tree finding software will be considered for these two complicated transmissions and general problems involved with the currently available approaches will be discussed. The main problems are rooting, contamination (especially between various recensions), and incomplete witnesses. Especially for the *Liber Aurelii* the available software proved not to be very helpful for the reconstruction of the stemmata.

Elisabet Göransson

Lund University / Centre for Theology and Religious Studies

A mix of stability and fluidity: An anthology or collection with variable content but rather stable text in a critical edition. Possible, recommendable?

Collections of sayings of the desert fathers and mothers are extant in manuscripts in many languages and are organized differently. The sayings were probably first written down in Greek during the fifth century, were translated into Latin, Syriac, Coptic, Palestinian Aramaic, Arabic and Ethiopic, Old Slavonic, Georgian and Sogdian already before the 9th century, and from the beginning of the 13th century into vernaculars all over Europe. They are “mixed-content miscellanies”: they include material that is variable both when it comes to appearance and order, but the text of each textual unit is rather stable.

Is a critical edition of one of these collections conceivable? The textual traditions of the collections of sayings are being studied in a large collaborative project in which philologists are working together and adding material to a relational database. The database and its interface is constructed as a combination of a relational database of the textual units with unique identifiers, and a xml/TEI- annotated corpus of transcriptions of the rich text traditions in many languages. In the relational database the unique identifier for each saying makes it possible to compare the collections over time, across different types of organisation, and interlingually.

The presentation focuses on the Latin text tradition as witnessed in the different collections: how it has been studied before, can be studied now, and possibly presented in a critical edition of the largest and most widespread collection that was first translated from Greek to Latin around 550 CE. After giving a background for the collections of sayings in Latin and for earlier editions and describing the kind of texts established in these editions, this paper will focus on the comparisons that can be made, the problems encountered and some options the editor may have. The function of a relational database in relation to the preliminary work towards a critical edition and the process involved will also be discussed.

Peter Robinson

University of Saskatchewan

Analyzing spellings across many manuscripts

The immense quantities of linguistic materials in medieval English vernacular manuscripts, broadly from the period between 1100 and 1500, offers at first glance a resource of vast promise for philologists. It appears self-evident that the existence of many hundreds of thousands of pages of medieval English texts, dating from throughout these four centuries and written all over England (with a significant number from Scotland, but not Wales and Ireland), must illuminate the development of varieties of English in this period. One might expect, for example, to see patterns of spellings in these pages distinctive of a particular time and space. Indeed, because every scribe had his or her own spelling, one might be able to identify the same scribe – or the same scribal group – at work in multiple manuscripts.

Many scholars, over many years, have responded to these opportunities. In medieval English studies, the most ambitious effort to analyze spellings across many manuscripts is the *Linguistic Atlas of Late Mediaeval English (LALME)*, originally a print publication and now available in electronic form (McIntosh, Samuels, and Benskin 1986; 2013). *LALME* sampled around 300 linguistic items across several thousand texts (each “text” being a single instance of writing by a single scribe), creating for each text a “Linguistic Profile”. Many surveyed texts can be dated and placed precisely, and *LALME* accordingly creates maps of the distribution of observed spellings across England. The *LALME* introduction cites impressive cases where manuscripts can be placed within ten to fifteen miles of a particular location, on the basis of the patterns of forms found within them. Other scholars have explored the spellings of particular scribes, or of specific sets of manuscripts, in search of the language of an author: thus the work of Jeremy Smith and Simon Horobin on Chaucer (Samuels and Smith 1988; Horobin 2007; 2003).

However, it has to be said that scholarly work in this area has been less productive than one might have expected. Consider for example just one case: the eighty-four manuscripts and four pre-1500 print editions of Geoffrey Chaucer’s *Canterbury Tales*, contained in some 29,000 pages of manuscript and print. Scholars have long remarked instances in this corpus of one scribe writing more than one manuscript: thus four of the earliest manuscripts appear to have been written by two scribes. Notably also, two of these four, both written by a scribe known as “hand d”, are stemmatically quite distant from each other. One would expect that analysis of the spelling forms of these two manuscripts, whether using *LALME* methodology or some quantitative tools, should show a clear linkage between the two manuscripts, and Jacob Thaisen explored this possibility in his De Montfort University doctoral thesis (Thaisen 2005).

Thaisen was limited in the data and tools available to him and was unable to establish any such linkage. In the years since, we have developed new bodies of spelling data for substantial sections of the *Tales*, and new analytic tools have become available. We will

make some of this data available before and at the conference, and report on attempts by ourselves and others to explore this data.

References

- Horobin, Simon. 2003. *The Language of the Chaucer Tradition*. Chaucer Studies 32. Cambridge,UK ; Rochester, NY: D.S. Brewer.
- . 2007. *Chaucer's Language*. Houndsmills, Basingstoke, Hampshire ; New York: PalgraveMacmillan.
- McIntosh, Angus, M. L. Samuels, and Michael Benskin. 1986. *A Linguistic Atlas of Late Mediaeval English*. [Aberdeen] ; New York, U.S.A: Aberdeen University Press.
- McIntosh, Angus, M.L. Samuels, and Michael Benskin. 2013. "A Linguistic Atlas of Late Mediaeval English." 2013. <http://www.lel.ed.ac.uk/ihd/elalme/elalme.html>.
- Samuels, M. L., and J. J. Smith. 1988. *The English of Chaucer and His Contemporaries: Essays*. Aberdeen: Aberdeen University Press.
- Thaisen, Jacob. 2005. "Studies in the Orthography and Some Manuscripts of Chaucer's "Canterbury Tales." PhD Thesis, Leicester, UK: De Montfort University.

Michael Stolz

University of Bern / Institute of Germanic Languages and Literatures

The Stemmatalogy of Reading Practices: Reconstructing the Library of Sigmund Gossembrot, a Southern German Humanist of the 15th Century

Sigmund Gossembrot (1417–1493), active in the period of early German humanism, has left a remarkable collection of manuscripts (containing classical authors, medieval texts on religious and secular matters, as well as contemporaneous literature) – a corpus that so far has only partially been explored. After completing a busy career as a civil servant in Augsburg, Gossembrot moved to Strasbourg where he joined the convent of the Knights Hospitaller *Zum Grünen Worth* (the 'Green Isle') for studying his books in the context of local libraries. Today, the surviving volumes are spread in archives all over Europe (with a certain concentration on the Bavarian State Library in Munich). Gossembrot left numerous annotations on the pages that attest his manifold literary interests and his reading habits, embedded in the social environment of both imperial towns and beyond. The abundant glosses, including many cross-references pointing to (owned and external) manuscripts with similar topics, also allow for the reconstruction of currently lost codices and their content. This paper discusses methods of documenting and examining Gossembrot's library, including a digital database currently under construction on: www.gossembrot.unibe.ch. A special focus will lie on the question, if Gossembrot's way of interconnecting manuscripts by cross-references can be analyzed in terms of stemmatological thinking and methods: A 'stemmatology' of Gossembrot's reading practices would include the (bi)directional relationships of his manuscript referencing, and the emergence of 'knots' in his literary interests.

References

- Joachimsohn, Paul: Aus der Bibliothek Sigismund Gossebrots, in: *Centralblatt für Bibliothekswesen* 11 (1894), p. 249–268, 297–307.
- Stolz, Michael: *Otium et negotium*. Reading processes in Early Italian and German Humanism, in: *Reading Books and Prints as Cultural Objects*, hg. von Evangelina Stead, Cham: Bodell & Brewer, 2018 (= New Directions in Book History), p. 81–106.
- Stolz, Michael: Transversale Lektüren. Die Bibliothek des Frühhumanisten Sigmund Gossembrot, in: *Die Bibliothek – The Library – La Bibliothèque. Denkräume und Wissensordnungen*, ed. by Andreas Speer and Lars Reuke, Berlin/Boston: de Gruyter, 2020, p. 484–507.

Jamie Tehrani and Gessica Martini

Durham University, UK / Department of Anthropology

Cinderella's Family Tree

Folkloristics, Phylomemetics and Population Memetics

The term “phylomemetics” was coined by Christopher Howe and Heather Windram to refer to the use of quantitative biological phylogenetic techniques to reconstruct lineages of cultural transmission, from the common roots of related languages to the accumulation of modifications in hand-copied manuscripts. This paper aims to contribute to recent attempts to apply phylomemetic methods to oral traditions, where the aim is to trace the mutation and diversification of folk narratives as they get passed on from generation to generation and spread from society to society. Our study focuses on one of the most famous and wide-spread tales in the folktale record: Cinderella.

Thousands of Cinderella-like stories have been documented from around the world, which folklorists have attempted to classify into different “types” representing distinct, though related, international traditions. The most comprehensive of Cinderella typologies was developed by Anna Birgitta Rooth, who divided the tales into five principal types: A, B, AB, BI and C, and suggested several hypotheses pertaining to their origins and relationships to one another. Here, we test Rooth's theories on a sample of 266 versions of Cinderella using Bayesian phylogenetic inference, phylogenetic networks (Neighbor-Net) and a model-based clustering method that was originally designed to elicit population structure from multi-locus genotype data (implemented in the program STRUCTURE).

While our results found some support for Rooth's typology, they indicate that the traditions are more analogous to demes than species in biology. “Interbreeding” among types appears to have been widespread, with one type (AB) revealed to be a hybrid of two older types (A and B), rather than a transitional form (A -> AB -> B), as had been previously suggested. While the extent of reticulate evolution greatly complicated the Bayesian and even NeighborNet analyses, the STRUCTURE analysis demonstrated that it was still possible to delineate and quantify the influence of distinct ancestral sources on the variation observed in contemporary versions of Cinderella. Our study therefore illustrates the potential value of incorporating a “population memetic” approach to the phylomemetic tool-kit, especially when dealing with highly contaminated datasets.

Tuomas Heikkilä

University of Helsinki / Faculty of Theology, Church History

Stemmatology, oral literacy, and folklore: results, ideas, and caveats

This paper aims to explore the possibilities of studying orally transmitted folklore applying computer-assisted phylomemetic approaches. Does the fluid nature of folklore allow using methods developed for stemmatological study? Is the “evolution through mutation” model valid for folklore studies?

Oral and written texts have coexisted in a constant interplay for thousands of years. In the modern academia, however, they are often being studied as separate entities by different disciplines. Still, a closer look at the methodology of textual and folklore scholars – or at the history of the two disciplines, for that matter – reveal several points in common. Just as the traditional textual criticism was built upon the idea of reconstructing the flawless original version of a text based on the existing witnesses, the so-called historical-geographic method of folklore studies aimed at finding the “masterful” original form of orally-transmitted folklore. Both disciplines shared a methodology based on a close comparison of several versions, and both often resulted in idealized heritage-objects.

The basic ideas, methods, and results of both textual criticism and historical-geographic method of folklore studies have been severely questioned during the past decades. In my view, both disciplines and their study objects share so many aspects that they would benefit from a closer cooperation. On one hand, it is obvious that many texts – e.g., hagiographical legends, fairytales etc. – used folklore as their sources. On the other hand, folklore is known to be influenced by texts.

I will showcase the traditional challenges of the historical-geographic method of folklore studies and the new possibilities of applying computer-assisted stemmatological methods in the context of the study of the orally transmitted tradition of *Death-Psalm of Bishop Henry* (Fi. *Piispa Henrikin surmavirsi*). The application of phylomemetic methods on folklore is certainly promising, but the question remains what the folklore studies might have to offer for the study of textual traditions.

SHORT PRESENTATION

Jost Gippert

University of Hamburg / Centre for the Study of Manuscript Cultures

Branching in Early Bible transmission (the DeLiCate project)

Under the title “The Development of Literacy in the Caucasian Territories” (“DeLiCaTe”), a new research project (ERC) at the Centre for the Study of Manuscript Cultures in Hamburg investigates, among other things, the interrelationship of translated texts in Armenian, Georgian, and Caucasian Albanian from the early centuries of their literacy (ca. 5th–10thcenturies). On the example of Biblical passages from the New Testament, the present paper illustrates peculiar characteristics of certain branches of the tradition and their dependency on divergent witnesses of the Greek and Syriac Bibles, showing that the split into branches must have occurred fairly early, manifesting itself in stemmatological clusters across several languages.

DIRECTED DISCUSSION

Aidan Conti || University of Bergen

Translation and composite texts

Building on topics broached by Caroline Macé's paper on the top of the stemma, I propose a discussion of and will pose questions on translations and composite texts. I will draw on a few specific examples from the Latin translations of sermones 15 and 17 (CPG 5524 and CPG 5526) of pseudo-Eusebius Alexandrinus. These homilies could and did circulate as two independent pieces and/or as one composite piece in both Greek and Latin. Questions that arise concern the identification of and stemmatological implications of 'faulty' translation, and the stemmatological relations between individual and composite texts.

Marina Buzzoni || Ca' Foscari University of Venice

Normalisation

In traditional stemmatology a distinction is primarily to be made between substantial readings and formal ones: usually, only the former are clues for determining the genealogical relationships between witnesses. Computer-assisted quantitative methods, however, take into account all variation – including, for example, spelling differences, abbreviation marks, and different letter forms – that may later undergo a process of normalisation for the purposes of stemmatic analysis. The extent of normalisation depends on the scholars' judgement and the methods they adopt. How far do the judgement and the methods adopted to normalise the input affect the output in digital stemmatology? Do they lead to an overestimation or an underestimation of the plesiomorphic vs apomorphic textual characters? The answers to these crucial questions will be illustrated by paradigmatic example.

Odd Einar Haugen || University of Bergen

Transliteration

When transcribing texts in the Latin alphabet, a major consideration is to decide when a glyph is no more than a variant and when it should be regarded as a character in its own right. A minimal pair test is typically used, so that if there is no difference in meaning when one glyph is exchanged for another, they are deemed to be variants of a single character rather than different characters. In the transliteration of a transcription, however, this test is far more uncertain, and while it may be possible to establish hard and fast rules for the transliteration from one script to another, it is not always straight forward to perform this operation in the opposite direction, i.e. from the transliterated text to the original one. This simple, but fundamental problem will be illustrated by the transliteration of Runic script to the Latin alphabet and vice versa, and implications of this problem will be considered.