Do Visual Features Matter? Studies in Phylogenetic Analysis of Mensural Music

Anna Plaksin
Max Weber Stiftung | Birmingham City University
plaksin@maxweberstiftung.de

Abstract

This paper reports on the task of developing concepts for a computational analysis of the transmission of mensural music based on concepts of phylogenetic analysis. Since the analysis of transmission aims for the reconstruction of relations between sources, it focuses on the differences of rather similar items. Therefore, it is necessary to find substitution models which are optimized for distinguishing fine levels of differences and to deal with the structural ambiguities and visual variance of mensural notation.

Introduction

The model of semantic domains in music notation is not only well known in the field of music encoding but is used as a common ground in reasoning about the representation of notated music. By modelling these domains in separate attribute classes, MEI provides a powerful feature offering the possibility to depict complexities of different kinds of music notation, e.g. mensural notation. Especially the lack of stable relationship between symbols and their interpretation is easily observed when encoding mensural music, but stemmatic analysis is typically led by the concept of significance commonly embodied in focussing on substantial variants, variants in pitch and duration. However, in the case of mensural music sources, with their richness of visual variance, where the particular context affects the process of reading and deciphering as well as developments in the notational system and varying concepts in mensural theory, that distinction reaches its limits:

> "The extent to which these can be considered 'non-substantive' is questionable: the positioning of line breaks, for instance, will have an effect on an editor's interpretation of the duration of manuscript accidentals, or stem direction may actually have an effect on rhythm in certain notational styles (as in some brands of 14th-century notation)" [1, p. 143].

This paper reports on the task of developing concepts for a computational analysis of the transmission of mensural music based on concepts of phylogenetic analysis. Starting with encodings of the sources of Josquin's *Missa D'ung aultre amer* and *Tu solus qui facis mirabilia*, it raises the question, how picking properties of mensural notation affects the resulting tree.

What to compare?

One main concern in methodological design is the preservation of the research object throughout the analysis. When dealing with questions, for example, like the effects of stress on people, the first and foremost task is the quantification of stress -- how can something so vague be detected by the means of measurable qualities. And even though we're dealing with music this question still matters. The analysis of the transmission of mensural music is a task that is inherently focused on the witnesses of this tradition, the sources itself. But these sources aren't digital objects. To make them available to such a task, the digitization the information about them is an inevitable bottleneck. And in this regard representation becomes crucial: If the representation of the source becomes distorted during the process of digitization, the whole analysis becomes flawed.

But, what in that particular context is the source and what features of it need to be maintained? First of all, when dealing with the transmission of music, luckily features of physical substance of the object could be regarded as subordinated. We're not interested in the object itself but in its role as a witness of the human

interaction with it, which is writing down music notation. And by following that track, because the sources are the only remaining witnesses of this interaction, we are faced with all the particularities and lapses that come with it.

First of all, a source of a piece of music isn't just the piece of music. A source can be erroneous, even to a point where the performance of the piece of music it bears isn't possible anymore.[1] Ambiguities in the position of notes on a staff are as likely as the challenge of deciding how long that line which represents a rest actually is.

And deciphering the notation is a complex task in itself. Limiting the scope to mensural notation, one main idiosyncrasy is the ambiguous relationship between a sign, its meaning and its result on performance. For this reason, visual variance of notation is not just likely but typical.[2] Hereby, the coexistence of single note shapes and ligatures adds to the various possibilities, but with only limited complexity.

For example, the phenomenon known as color minor is a one of these special cases. Usually, in transcription to modern Common Western Music Notation it is treated equally to the *punctum augmentationis*, even if they exist commonly in close succession (see e.g. figure 1) [3, p. 138f.]. But why would both manners be used side by side? While Stanley Boorman [4, pp. 72-75] describes possibilities of further implications, Ronald Woodley [5] describes a change in notation practice around 1500. Which position is to be followed doesn't really matter in this context -- rather the range of subtleties has to be considered.

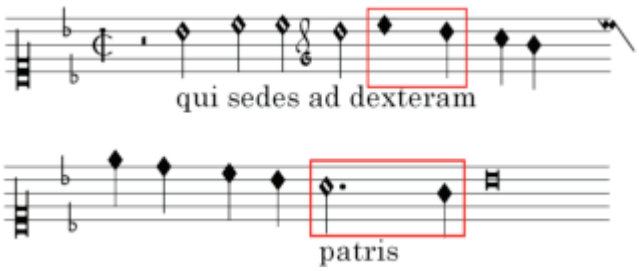


Figure 1: *Color minor* and punctuation is used in close succession. M. D'ung aultre amer, Gloria, Superius, after [VatS 41, fol. 150v].

As well, a focal point for the transmission of the *M. D'ung aultre amer* and *Tu solus qui facis mirabilia* are varying signs for the *sesquialtera* (see figure 2). It is not just a layer of visual variation but could be seen as well as a symptom of changes in the practice of musical notation [6]. As Anna Maria Busse Berger explains, the understanding of proportions evolved from a "substitute for mensuration signs" [7, p. 185] to a self-contained sign for diminution [7, pp. 182-96]. The variant reading of a *sesquialtera* including a change to *tempus perfectum* with a circle added on top of the **3** might be understood as effect of this trend. Following this argumentation, it would be essential to track this kind of variants during the whole process.

Figure 2: Different signs for sesquialtera: While (a) give no statement about the mensuration, (b) signals perfect time unambiguously [7, p. 230].

Bringing these aspects together, there are some valid conclusions concerning the machine-readable representation. When sources are the objects of our particular interest, it is necessary to allow for their idiosyncrasies to be pivotal, both unperformable corruptions in the source and the fine subtleties of notation. With the former, even if it means to accept that ‚the music' cannot be read from a manuscript [8, p. 169]. But when reading the music is impossible, a source still has its documentary value that can be captured. And moreover, con-

1 Like e.g. a suddenly ending superius in the Osanna of Dufays *Missa Se la face ay pale* in [VatS 14, fol. 35v].

2 "Mensural notation is at least partly redundant in that the scribe often has a choice of representing a certain musical content in different visual manifestations" [2, p. 58].

cerning notational idiosyncrasies, it is also possible to trace them from a point of describing the notation itself, without trying to resolve durations, reconstructing a conceptual piece of music – meant as thinking the parts together – or even suggest appropriate inflections of pitches. But the question arising from these thoughts is: Does that actually work in a computational analysis of transmission when similarity is usually estimated based on the perception of this conceptional or the aural dimension of music?

Distinction of difference

Other endeavours to use global sequence alignment for notated music [9], [10], and [11] focus on either retrieval scenarios or minimizing differences of notation, mode and/or tempo. Analysing patterns of transmission comes with a different scope. The goal of stemmatic analysis is more or less giving statements about the relationship between sources based on their variants. This means, instead of querying the most similar in a heterogeneous group, the main task is to cluster a group of rather similar objects according to their differences. To allow this clustering, we might not focus on their similarities but rather to distinguish the degree of deviation between a group of sources. Therefore, it is necessary to find substitution models which are optimized for distinguishing these fine levels of difference. But how could these models be developed without any advanced experience in measuring difference of mensural music?

On the one hand, there are mathematical models: But they are focused on similarity scores used in local sequence alignment. And they need to follow particular assumptions to be valid. The expected similarity score of an alignment of random sequences needs to be below zero while there is at least one positive score. Conversely this means it would be necessary to decide what level of similarity is to be denominated as neutral similarity: S = 0 What is different enough to be similar but similar enough to not be different?

On the other hand, there are already existing stemmata, made very cautiously e.g. while editing a certain piece of music. But analysing these shows, that every stemma is constructed strictly on its own terms. When comparing the stemmata of the joint transmission of Josquins *Missa D'ung aultre amer* and the motet *Tu solus qui facis mirabilia* – which is used as a replacement for the Benedictus and Osanna II – crucial disagreements become evident. And because of these disagreements, pre-existing stemmata cannot serve as a benchmark as well.

An obvious conclusion in addressing these challenges is using methods with few external prerequisites. First of all, a global alignment using distance-based substitution models is the chosen approach. By stating that identity as distance D = 0 the definition of neutral similarity is avoided. Moreover, a data-based process was developed for evaluating substitution models. Based on the method of surrogate data analysis [12], an approach was chosen, that scales the strength of separating levels of distance. Hereby sequences are shuffled to provide independent and identically distributed random sequences as a benchmark.

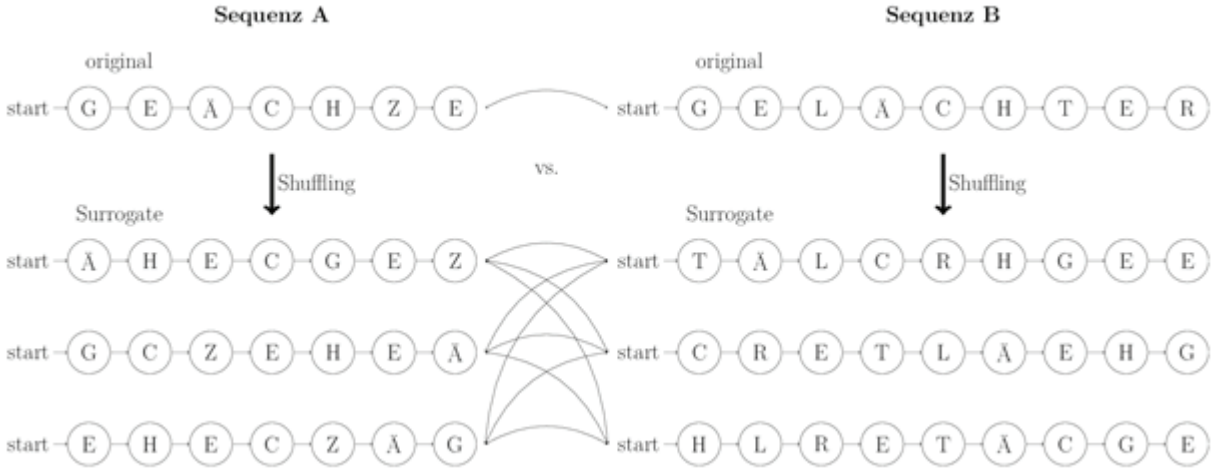


Figure 3: Figure 3: Comparing original data against independent and identically distributed random sequences.

This analysis makes use of three main preconditions:
First of all, its use in the course of finding models for the analysis of transmission depends on dissimilarity as the central criterion of stemmatic analysis. Second, the shuffling utilizes the assumption that the internal structure of a sequence is constitutive for similarity, in the way that the order of letters constitutes a word. And the third condition is, that the relative distance between original and surrogate comparisons is affected by the similarity of the original sequences. Therefore, it must be possible to estimate the dissimilarity of the original sequences by quantifying the deviance of the relative distance between these two original sequences and their shuffled surrogates. Moreover, this approach serves as the basis for an analysis of variance to detect a trend in comparing sequences, grouped by an estimated level of similarity. And observing the behaviour of a set of attributes with this test set-up can lead to an informed choice of analytical parameters.

Comparing feature sets

In this analysis of variance, the joint *M. D'ung aultre amer* / *Tu solus transmission* serves as the test case.[3] To detect a trend in the deviance of relative distance between original and surrogate data, groups of estimated similarity has been defined by the non-controversial relations of those two conflicting stemmata, together with arbitrarily chosen groups:

1. Different piece of music: Quis dabit capiti meo aquam
2. Different parts of the same mass section from the same source
3. Different parts of the same section from the same source
4. *Tu solus*: Mass vs. motet tradition
5. *Tu solus*: Same tradition
6. *Tu solus*: Direct dependency
7. Same part before and after scribal intervention

Since the main question that arose during the encoding of the sources is how much interpretative encoding of mensural notation is least necessary for performing, the tested parameter sets are mainly designed to capture certain states of interpretation of mensural notation.
The first state, labelled as **signbased.vis** is similar to recognising and describing symbols. Every symbol in a staff is described independently depending on the kind of symbol, mostly based on the element names and attributes used in the encoding. Regarding mensuration signs and proportion signs, only an identifier classifying the visual sign is used for further discrimination. The written pitch, which is used as a feature for notes and accidentals, can in this regard be seen as a classification of the vertical orientation within a staff – no further inflection of accidentals or *musica ficta* is intended. As well, notes and rests are merely distinguished by their types as encoded with *@dur*. In addition, notes have features regarding their coloration and the form of a ligature and their position within a ligature.

In contrast, **signbased.log** still records every symbol in the staff, but tries to capture the actual meaning of the symbols. This means, for mensuration and proportion signs, it records the Tempus, Prolatio, Modus minor and maior and the *@num/@numbase*. As well, the duration of notes and rests are resolved into relative durations, and the pitch is recorded including resolved written accidentals.

Since the aim of comparing sources makes it inevitable to follow one source as it presents itself without emendation of errors, a parameter set containing performance-related information would undermine this. Already the observance of written accidentals is a grey area, but resolving *musica ficta* is in this regard too interpretative a task. Therefore, data that needs to have taken more than a single part into consideration is excluded. Instead, another parameter set has been created as a further reduction to substantial parameters. The parameter set called **superlogical.gap** takes only notes and rests with their relative duration and resolved pitch into consideration. In this way, it mimics focussing exclusively on substantial variants. And in addition, another parameter set **signbased.all.gap** contains every created parameter.

3 For a more detailed description of the analysis, see [13].

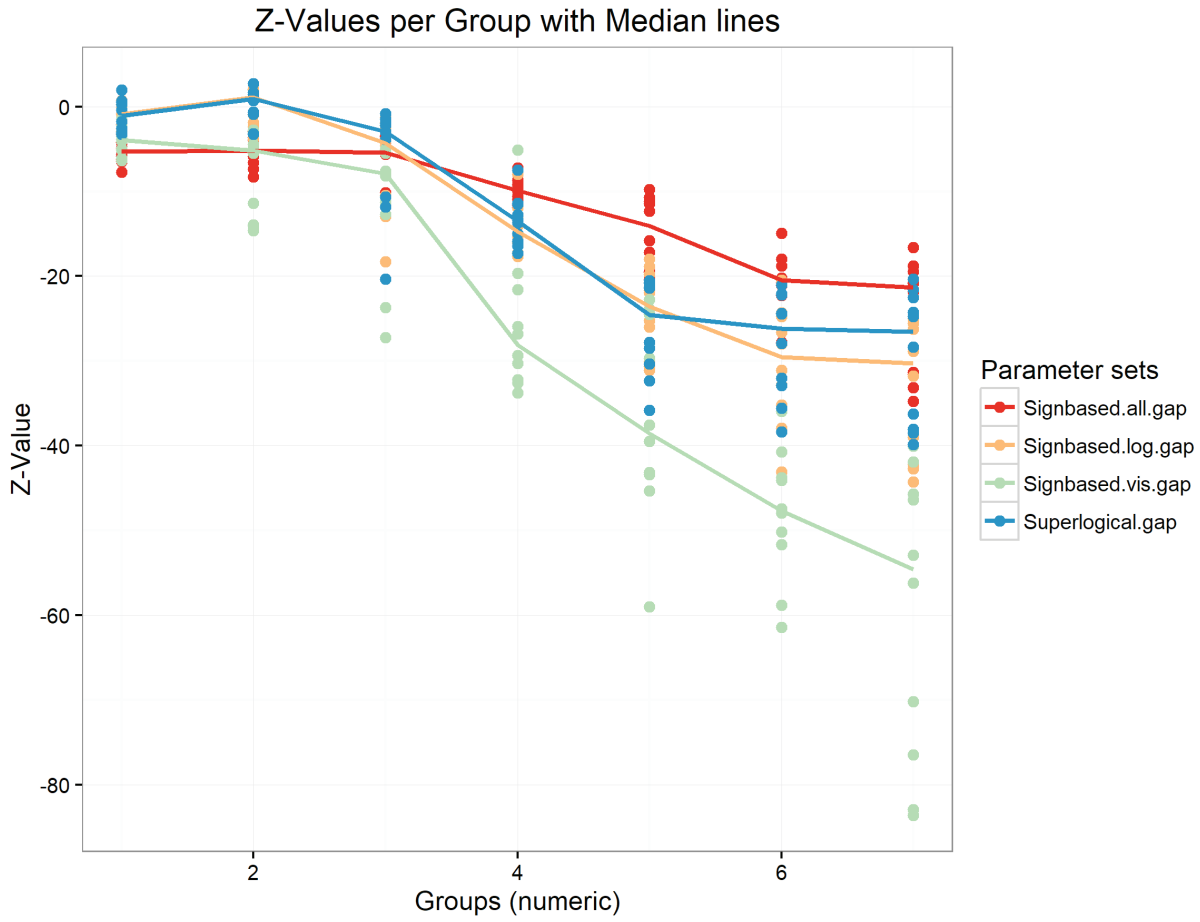


Figure 4: Deviance of distance between original and surrogate sequences per parameter set and group (see list). Median lines visualize trend.

By performing the comparison of original and surrogate sequences per parameter set and per group, several crucial observances can be made (see figure 4). The **signbased.all.gap** set not only have the slightest slope, but apart from that, the deviance between original and surrogate is already significant for the comparison of different pieces.[4] Having this group as a control group, which sets another piece of music at random against the chosen example, an acceptable parameter set must show now significant deviance from random comparisons, whereas the other sets match this demand. In conclusion, a set containing all parameters is not appropriate at all for this task.

The parameter sets **signbased.log.gap** and **superlogical.gap**, mainly focussing on the logical meaning of a symbol in the context of notation or the resulting impact as a series of notes and rests with a certain relative duration and a pitch, show a very similar behaviour. There is a visible difference between the arbitrarily chosen groups and the quite similar groups 5-7, with the group comparing different traditions in the middle. But the differences between the similar groups are hardly distinguishable. And moreover, the minimal deviance between original and surrogate is not the control group but the group comparing different voices. This could be explained by a high influence of absolute pitch. These parameters, therefore, might serve well in a setting of retrieving similar pieces from a heterogeneous group.

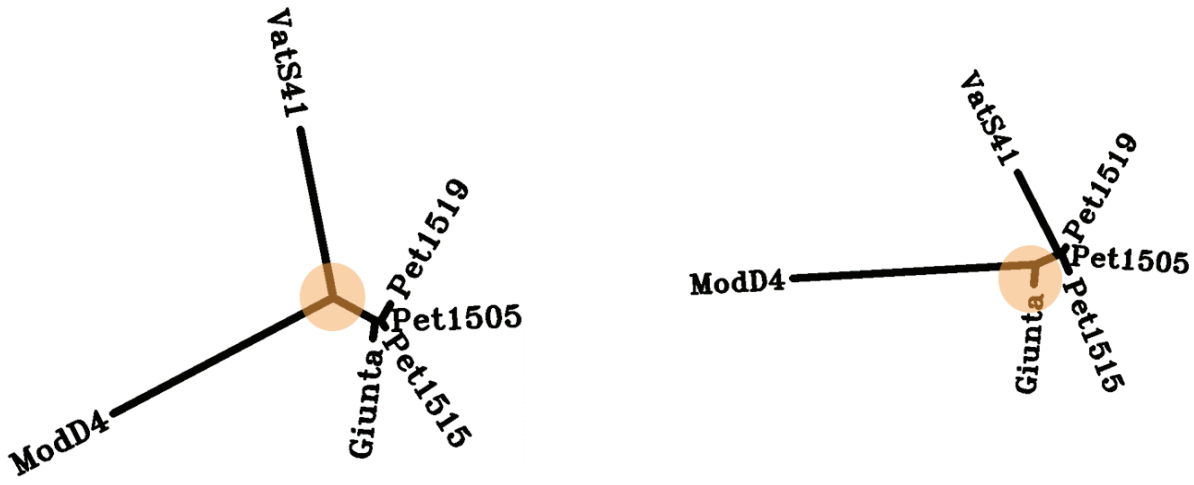
But when analysing transmission, the task is to cluster similar pieces according to their differences. In this regard, the set **signbased.vis.gap** seems more appropriate. It distinguishes well between the arbitary groups and the ‚realistic' groups and its minimum is at the control group. Moreover, it shows the steepest slope for the groups 4-7 based on realistic comparisons.

4 Using a Wilcoxon signed-rank test (α= 0.001): W = 17851.

Comparing trees

But beside the results of the surrogate data analysis, it is worth to take another look at trees. As already mentioned, the transmission of the *M. D'ung aultre amer* and *Tu solus qui facis mirabilia* is of relevance because there are two conflicting stemmata [14, p. 34], [15, p. 43]. In detail, they show how much weighing different aspects and focussing on certain variants can lead to different points of view [13, pp. 79-82]. Crucial for the diverging layout is whether proportion signs are taken into consideration or not. While Noblitt ignores them, Blackburn traces one of the *sesquialtera* signs back to Petrucci' editor Petrus Castellaus [15, p. 40]. Conversely, the other variant is treated as authorial.[5] At the first glance, this might fit questions this paper is stating when observing the effect of a single visual feature on a stemma. But this is covered by another aspects. The central criterion for a stemma is usually significance, the likelihood of the same error occurring independently. The question of notational parameters is on side of the question, the other is weighing the influence – which has been done by stating authorial influence.

Therefore, it might be rather useful to start with a delimitation. The trees constructed as a part of this study are unrooted trees based on global distances of sequences.[6] Without root, they give no hint about a possible origin and in addition no information about relations to that origin. And, in contrast to a stemma, these trees don't follow any rule of significance. Every dissimilarity that has been detected by the chosen parameter set affects the tree.



(a) Signbased.log.gap.

(b) Superlogical.gap

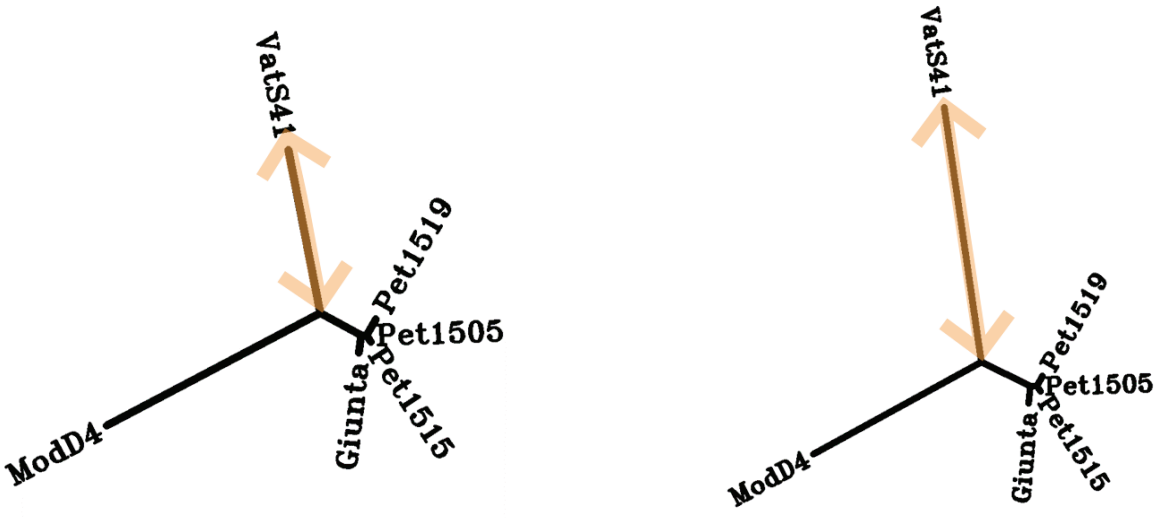
Figure 5: .Instability of unrooted trees: Superlogical.gap leads to a different layout than signbased.log.gap, disregarding that [Gio1526] is a reprint of the Petrucci prints.

Figure 5 shows the trees regarding the superius and tenor of the mass cycle[7] constructed on the basis of the *superlogical* and the *signbased logical* parameters. Obviously, the topology of both trees differs. While the printed sources are grouped together in figure 5a apart from the two manuscripts [ModD 4] and [VatS 41], the superlogical tree (figure 5b) sets only [ModD 4] apart. Moreover, it puts the Vatican source in closer relation to the three Petrucci prints than the Giunta reprint of 1526 – a highly doubtful result. In this regard, the tree based on the parameters per sign fits better to external knowledge about the sources.

5 "Josquin himself normally used '3' to indicate sesquialtera" [15, p. 40]. As well Gaffurius attests to Josquin's use of "3" [6, p. 418f.].

6 The neighbor-joining algorithm according to [16] has been used.

7 Only superius and tenor were available in all sources.



(a) Signbased.log.gap. (b) Signbased.vis.gap

Figure 6: .Influence of ligatures and coloration: Visual parameters affect the distance between the Cappella Sistina choir book and other mass sources.

In comparison, figure 6 compares the trees built on parameter sets either of uninterpreted notation and of signbased resolved meanings. Obviously, the layout of both trees is identical, while the edge lengths differ. In particular, the distance of the Vatican source in relation has grown. When taking a look at the sources, this result is evident. While [ModD 4] conforms mostly with the printed sources regarding the use of coloration and ligatures, the Cappella Sistina choir book makes a heavy use of coloration and ligatures, in melismatic sections of the tenor not unlikely complex multi-note ligatures. Notably, the variant *sesquialtera* signs don't change the layout of the tree, otherwise there would have been a difference between the trees derived from the signbased models – only **signbased.vis** takes the sign into account together with ligatures and coloration. Instead, it weighs much more if the model is signbased or notation agnostic.

Conclusion

First of all, I would like to start concluding about the analysis of transmission. The presented experiment shows a different usage of sequence alignment than retrieval scenarios. When reusing methods of phylogenetic analysis, the focus lies on the distinction of difference rather than on finding similarities. Therefore, other models need to be used. As well, it is obvious, that an unrooted tree constructed on global alignments must be read differently than a stemma because it is based on other conditions. But when trees are constructed based on different feature sets, the effect of certain assumptions can be made visible. Whereas a stemma usually relies on few significant variants, the showed trees take all detected differences into account.

Moreover, the results can clearly be summarized: Notation matters! The presented method of surrogate data analysis gives hints about the specific behaviour of a model. It shows, for the purpose of discriminating rather similar items, the model based on uninterpreted mensural notation provides better results than the models using resolved durations and ignoring visual variance. And while the surrogate data analysis shows no big difference for the latter, the comparison of trees favours a notation specific approach. These results not only demonstrate that it is possible to bypass the interpretative reading of music notation for the purpose of comparing differences, but as well it makes clear that notation itself can be a subject of research, e.g. in tracing changes in notational praxis. For this research, encoding is a pivotal part, but under different circumstances than e.g. for machine-readable editions or musical analysis. Encoding, used as a structured description of notation and its idiosyncrasies can serve the analysis of notation and its interpretation.

With that purpose in mind, the separation of semantic domains as provided in with MEI is a powerful tool. For the matter of encoding mensural notation, this is a highly intricate task, since it means to differentiate carefully the levels of interpretation. In the case of tracing the ambiguous relationship of sign and meaning, a procedural approach would be favourable, classifying the sign, representing its instructional value and illustrate resulting consequences in performance.

List of Sources

[Gio1526] Jacopo Gionta, Giovanni Giacomo Pasoti, and Valerio Dorico, eds. Missarum Josquin Liber secundus. Rom, 1526. http://data.onb.ac.at/rec/AC09229566

[VatS 14] MS Capp.Sist.14. Biblioteca Apostolica Vaticana. Rome, Italy. https://digi.vatlib.it/mss/detail/Capp.Sist.14

[VatS 41] MS Capp.Sist.41. Biblioteca Apostolica Vaticana. Rome, Italy. https://digi.vatlib.it/mss/detail/Capp.Sist.41

[ModD 4] Ms. Mus. IV. Archivio Capitolare. Modena, Italy.

[Pet1505] Ottaviano Petrucci, ed. Missarum Josquin Liber secundus. Venedig, 1505. http://diglib.hab.de/drucke/2-8-musica-2s/start.htm

[Pet1515] Ottaviano Petrucci, ed. Missarum Josquin Liber secundus. Ave maris stella. Hercules dux ferrarie. Malheur me bat. La mi baudichon. Una musque de buscaya. Dung aultre amer. Fossombrone, 1515. http://data.onb.ac.at/rec/AC09319724

[Pet1519] Ottaviano Petrucci, ed. Missarum Josquin Liber secundus. Ave maris stella. Hercules dux ferrarie. Malheur me bat. La mi baudichon. Una musque de buscaya. Dung aultre amer. Fossombrone, [1519]; 1515. http://data.onb.ac.at/rec/AC09170773

Works Cited

[1] Dumitrescu, Theodor, and Marnix van Berchum. “The CMME Occo Codex Edition: Variants and Versions in Encoding and Interface” in *Digitale Edition zwischen Experiment und Standardisierung*, eds. Peter Stadler and Joachim Veit. Beihefte zu editio. Tubingen: Niemeyer, 2009, 129-46.

[2] Schmidt, Thomas. “Making Polyphonic Books in the late Fifteenth and early Sixteenth Centuries” in *The Production and Reading of Music Sources*, eds. Thomas Schmidt and Christian Thomas Leitmeir (Epitome Musical). Turnhout: Brepols Publishers, 2018, 1-98.

[3] Apel, Willi. *Die Notation der polyphonen Musik: 900 - 1600.* 5. Au. Wiesbaden: Breitkopf & Härtel, 2006.

[4] Boorman, Stanley. “Notational Spelling and Scribal Habit”, in *Datierung und Filiation von Musikhandschriften der Josquin-Zeit*, ed. Ludwig Finscher. Wolfenbütteler Forschungen. Wiesbaden: Harrassowitz, 1983, 65-109.

[5] Woodley, Ronald. “Minor coloration revisited: Okeghem’s Ma bouche rit and beyond” in *Théorie et analyse musicales*, ed. Anne-Emmanuelle Ceulemans (Publications d’histoire de l’art et d’archéologie de l’Université Catholique de Louvain. Musicologica neolovaniensia. Studia). Louvain-la-Neuve, 2001, 39-63.

[6] Blackburn, Bonnie J. “The Sign of Petrucci’s Editor” in *Venezia 1501: Petrucci e la stampa musicale*, eds. Giulio Cattin and Patrizia Dalla Vecchia (Serie III. Studi musicologici. B, Atti di convegni / 6). Venezia, 2005, 415-30.

[7] Busse Berger, Anna Maria. *Mensuration and proportion signs: Origins and evolution*. Oxford monographs on music. Oxford: Clarendon Press, 1993.

[8] Boorman, Stanley. “The Uses of Filiation in Early Music” Text 1 (1981), 167-84, http://www.jstor.org/stable/30234249

[9] van Kranenburg, Peter. “A Computational Approach to Content-Based Retrieval of Folk Song Melodies”. Utrecht University, 2010. PhD dissertation. http://hdl.handle.net/20.500.11755/8436a210-ceeb-4c66-ba02-ffe5c7e66a42

[10] Mongeau, Marcel, and David Sanko. “Comparison of Musical Sequences” Computers and the Humanities 24, no. 3 (1990), 161-75, http://www.jstor.org/stable/30200223

[11] van Nuss, Jelmer, Geert-Jan Giezeman, and Frans Wiering. “Searching musical incipits by means of sequence alignment” presented at the Music Encoding Conference, Tours, France, May 16-19, 2017.

[12] Theiler, James, Stephen Eubank, André Longtin, and Bryan Galdrikian. “Testing for nonlinearity in time series: the method of surrogate data” Physica D: Nonlinear Phenomena 58, nos. 1-4 (1992), 77-94. https://doi.org/10.1016/0167-2789(92)90102-S.

[13] Plaksin, Anna. “Modelle zur computergestützten Analyse von Überlieferungen der Mensuralmusik: Empirische Textforschung im Kontext phylogenetischer Verfahren“. Darmstadt: Technische Universität Darmstadt, 2019. PhD dissertation.

[14] Josquin. *Masses based on secular polyphonic songs: Critical commentary*, ed. Thomas Noblitt. New Josquin Edition 7. Utrecht: Vereniging voor Nederlandse Muziekgeschiedenis, 1997.

[15] Josquin. *Motets on non-biblical texts 2: Critical commentary*, ed. Bonnie J. Blackburn. New Josquin Edition 22. Utrecht: Vereniging voor Nederlandse Muziekgeschiedenis, 2003.

[16] Studier, James A., and Karl J. Keppler. “A note on the neighbor-joining algorithm of Saitou and Nei” Molecular biology and evolution 5, no. 6 (1988), 729-31. 10.1093/oxfordjournals.molbev.a040527.