Towards a Generative Model of Medieval Chant

Tim Eipert **Fabian C. Moss**
Julius-Maximilians-Universität Würzburg
Würzburg, Germany
{tim.eipert,fabian.moss}@uni-wuerzburg.de

1 Introduction

For some time now, musicologists studying premodern Europe have been examining the question of regional and contextual influences on vocal melodies (Huglo, 1999; Haggh, 1996; Kruckenberg, 2006; Dobszay, 2017). Since a large amount of digital data from different research projects is available (Dvořáková et al., 2026), there is significant potential for computational analysis. However, previous computational approaches to medieval melodies often focus on isolated features, like modality, or rely on purely descriptive statistics (Cornelissen et al., 2020; Helsen et al., 2023; Hiley, 2017; Weiss, 1964; Hughes, 2017; Eipert and Moss, 2023). While descriptive analyses can identify associations between variables, they cannot explain the underlying mechanisms by which these melodies were formed and transmitted. Here, we present initial steps towards developing a generative model for pitch sequences in monophonic medieval chants as functions of both musical and contextual parameters, learnable from data. Our main contribution is a model that simultaneously integrates multiple metadata covariates. Instead of providing a purely descriptive analysis or focusing solely on mode, we aim to model the data-generating process itself.

2 Model

We assume an overall model structure comprising several subcomponents, or submodels. The central part is a Markov model of order $n-1$ (also called an n-gram model; Manning and Schütze, 2003) over a pitch alphabet of size V. It consists of a starting pitch (or reference pitch) π and a matrix $\mathbf{A}$ that guides transitions to subsequent pitches. Both π and each row A_t of the transition matrix $\mathbf{A} \in \mathbb{R}^{T\times V}_{\geq 0}, |A_t| = 1$, are samples from a Dirichlet distribution. The starting pitch x_0 is drawn from a categorical distribution, parameterized by π. Each subsequent pitch $x_t, t = 1, \ldots, T$, again follows a categorical distribution, parameterized by A_{t-1}, where $t-1$ is the state of the previous pitch. The model is fully specified by

$$\begin{aligned} A_t &\sim \text{Dirichlet}(\alpha) \\ \pi &\sim \text{Dirichlet}(\beta) \\ x_0 &\sim \text{Categorical}(\pi) \\ x_{t+1} &= \text{Categorical}(A_t), \end{aligned}$$

with hyperparameters $\alpha = \beta := (1, \ldots, 1) = \mathbf{1}^V$.

To capture the complex interactions of extra-musical influences, we conceptualize our approach as a hierarchical generative model. In this framework, the specific contextual and musical entities act as distinct submodels. For instance, a regional submodel and a modality submodel each provide specific parameters that condition the base Markov model. A preliminary DAG of the causal paths of the parameters on the Markov model is shown in Figure 1. The categorical variables in the DAG are used as predictors, which in turn parameterize the hyperparameters α and β.

3 Musical content

Modality Modality is a concept that originated in the Middle Ages and was used to group structurally similar songs (Atkinson, 2009). In Figure 2, a chant characteristic for mode 7 (later called G-mixolydian) is shown. However, this notion of mode as structural similarity holds only for prototypical and not all examples. In later theoretical writings, the final note ('finalis') is often seen as decisive for the mode class (McAlpine, 2004). However, it has recently been shown that the remainder of a chant melody is also strongly influenced by mode, as chant embeddings segmented

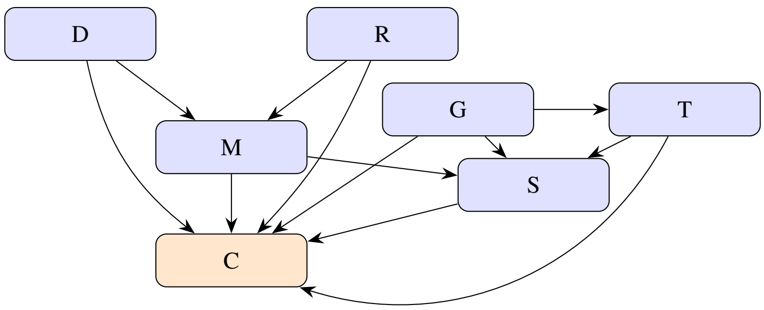

Figure 1: Example of a causal Directed Acyclic Graph (DAG) of the different influences on the melody: **D**ate and **R**egion influence **M**ode; **G**enre influences **T**ext; Genre and Text influence **S**egmentation. All of them have direct or indirect influence paths to the **C**hant melody.

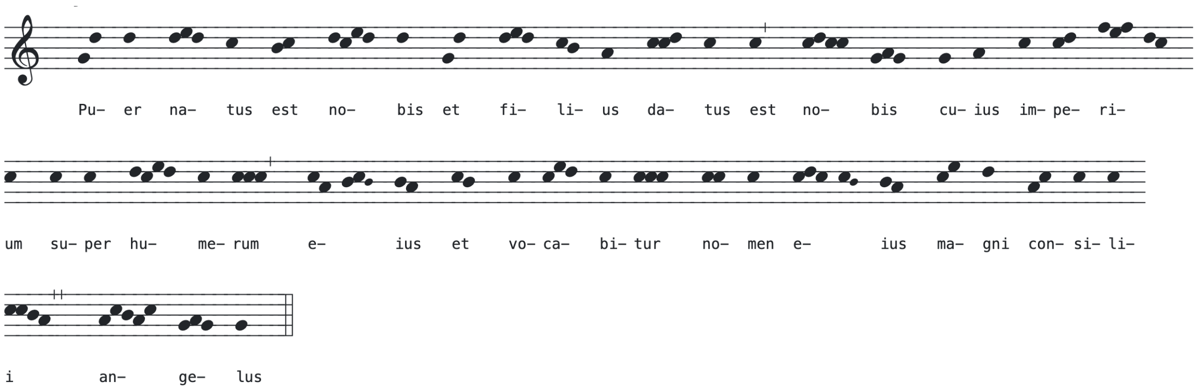

Figure 2: A transcription of a characteristic mode 7 chant. Source: `https://cantusdatabase.org/chant/638421`.

by n-grams yield better classification results than final pitch and ambitus only (Cornelissen et al., 2020). Nakamura et al. (2023) use a mixture Markov model in which multiple mode-dependent pitch transition matrices are inferred. This model is close to ours here, but is limited to mode alone as influence on the pitch sequence. One reason mode has such a pervasive influence on melody may be that it operates partly through a shared vocabulary of reusable melodic segments (Helsen et al., 2023). Lanz and Hajič jr (2025) provide support for this view by showing that an unsupervised, memory-optimal segmentation of chant melodies achieves state-of-the-art mode classification, suggesting that melodic segments carry mode-specific information.

Text In the Middle Ages, the lyrics of chants had a significant influence on melodies (Treitler, 2003, p. 17), e.g. they provide structure, as melodic endpoints coincide with semantic units of the lyrics. Additionally, different metrical structures of lyrics can influence melodic emphasis and thus the probability of pitch transitions. Finally, semantic interpretation of lyrics can also influence melodies.

Furthermore, the same text is often accompanied by different melodic variations in different manuscript traditions, reflecting the influence of geographical and temporal transmission on pitch sequences. (Eipert and Moss, in press). These variants can be systematically recovered through sequence alignment and phylogenetic inference, constituting another source of structured variation that our generative model aims to capture.

4 Contexts

Liturgical Function & Genre Most chants come from Christian liturgical sources, and they are to be performed in liturgical rituals. This influences melodies, e.g., their lengths and complexity (Hiley, 2009): antiphons, for instance, are structurally simpler and shorter than responsories, implying different distributions of pitch transitions.

Regionality The melodies' processes of transmission are complex: from Rome to the Frankish Empire, spreading further and further east, and then back to Italy (Planchart, 2017). These transmissions influence the melodic forms, partly due to transmission errors, because regional actors have an interest in leaving their own mark or because of the change of political borders (Eipert and Moss, 2026). This, alongside church regulations, resulted in interesting constellations in which a basic corpus of songs was handed down virtually unchanged, with only minor regional variations. Regional influence, therefore, depends on genre.

Date In most cases, we can only approximate the creation date of manuscripts by codicological analysis of their material. It is difficult to approximate the date of origin of the chants because the manuscript's age provides only an upper bound on the time a chant existed. Ballen et al. (2024) address this problem by applying Divergence Time Estimation (DTE) to a Bayesian phylogeny of chant melodies. With this method, they separate the evolutionary rate from time in the branch lengths of the inferred tree and assign probability distributions over dates to ancestral nodes. Chant manuscripts can be understood as incomplete snapshots of a musical culture at a specific point in time and place. To retrieve and organize this temporal and contextual metadata systematically across sources, we can rely on the Cantus Database (Lacoste, 2022), which provides the most extensive digital collection of Gregorian chant with source-level provenance and dating information.

5 Conclusion

Melodies of Gregorian chant are shaped by the rich interplay of musical and contextual factors that have so far only been studied in isolation. While deep learning architectures like the Transformer represent state-of-the-art techniques in sequence modelling, their black-box nature obscures the explicit musical mechanisms we aim to investigate. Bayesian methods allow us to preserve interpretable causal structures. In our ongoing research, we propose a hierarchical generative model that integrates these influencing factors–modality, text, segmentation, ritual function, regionality, and date–as explicit covariates to moderate a Markov model of pitch transitions. Using this model, we can reveal causal structures that purely descriptive methods cannot access. For example, by conditionalizing the text, the model can isolate residual regional melodic variations that cannot be explained only by textual structure. Similarly, the model can reveal whether melodic fragments of a particular pattern are stable in different regional traditions, or whether parts of what appear to be pattern vocabulary are regional products, which only becomes apparent when both covariates are included.

References

Charles M. Atkinson. 2009. *The Critical Nexus: Tone-System, Mode, and Notation in Early Medieval Music*. Oxford University Press.

Gustavo A. Ballen, Klara Hedvika Muhlova, and Jan Hajic. 2024. What did the dove sing to Pope Gregory? Ancestral melody reconstruction in Gregorian chant using Bayesian phylogenetics. https://doi.org/10.1101/2024.08.19.608653.

Bas Cornelissen, Willem H Zuidema, John Ashley Burgoyne, et al. 2020. Mode classification and natural units in plainchant. In *ISMIR*. pages 869–875.

László Dobszay. 2017. Concerning a chronology for chant. In *Western Plainchant in the First Millennium*, Routledge, pages 217–229.

Anna Dvořáková, Tim Eipert, Debra Lacoste, and Jan Hajič Jr. 2026. Making chant computing easy: Cantuscorpus v1. 0 and the pycantus library. *Transactions of the International Society for Music Information Retrieval* 9(1):164–178.

Tim Eipert and Fabian C. Moss. 2023. MonodiKit: A data model and toolkit for medieval monophonic chant. In Martha E Thomae, editor, *Proceedings of the 10th International Conference on Digital Libraries for Musicology*. Association for Computing Machinery, New York, NY, USA, DLfM '23, pages 67–71. https://doi.org/10.1145/3625135.3625145.

Tim Eipert and Fabian C Moss. 2026. Inferring Communities of Medieval Music Manuscripts using Stochastic Block Models. *Transactions of the International Society for Music Information Retrieval* 9(1):36–49. https://doi.org/10.5334/tismir.298.

Tim Eipert and Fabian C Moss. in press. From JPEG to Tree: Analysis of Melodic Variants in the Graduale Synopticum Edition. In Elsa De Luca and Martha Thomae Elias, editors, *Shaping Early Music Research in the Digital Era: Case Studies on Plainchant*, Brepols.

Barbara Haggh. 1996. Mode in late-medieval plainchant from Cambrai. In *Modality in the Music of the Fourteenth and Fifteenth Centuries/Modalität in*

Der Musik Des 14. Und 15. Jahrhunderts, Hänssler-Verlag, Neuhausen-Stuttgart, Musicological Studies and Documents, pages 129–147.

Kate Helsen, Mark Daley, and Jake Schindler. 2023. The sticky riff: Quantifying the melodic identities of medieval modes. *Empirical Musicology Review* 16(2).

David Hiley. 2009. *Gregorian chant*. Cambridge University Press.

David Hiley. 2017. Some observations on the interrelationships between trope repertories. In *Embellishing the Liturgy*, Routledge, pages 25–33.

David G Hughes. 2017. Further notes on the grouping of the aquitanian tropers. In *Embellishing the Liturgy*, Routledge, pages 275–284.

Michel Huglo. 1999. Division de la tradition monodique en deux groupes "est" et "ouest". *Revue de musicologie* pages 5–28.

Lori Kruckenberg. 2006. The lotharingian axis and monastic reforms: Towards the recovery of an early messine trope tradition. In *Cantus Planus–Study Group of the International Musicological Society: Papers Read at the Twelfth Meeting, Lillafüred, Hungary. 23–28 August 2004*. pages 723–752.

Debra Lacoste. 2022. The Cantus Database and Cantus Index Network. In Daniel Shanahan, John Ashley Burgoyne, and Ian Quinn, editors, *The Oxford Handbook of Music and Corpus Studies*, Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190945442.013.18.

Vojtěch Lanz and Jan Hajič jr. 2025. Gregorian melody, modality, and memory: Segmenting chant with Bayesian nonparametrics. https://doi.org/10.48550/arXiv.2507.00380.

CD Manning and H Schütze. 2003. *Foundations of Statistical Natural Language Processing*. MIT Press, 6 edition.

Fiona McAlpine. 2004. Beginnings and endings: Defining the mode in a medieval chant. *Studia Musicologica Academiae Scientiarum Hungaricae* 45(1-2):165–178.

Eita Nakamura, Tim Eipert, and Fabian C Moss. 2023. Historical Changes of Modes and their Substructure Modeled as Pitch Distributions in Plainchant from the 1100s to the 1500s. In Kitahara, Tetsuro, Aramaki, Mitsuko, Kronland-Martinet, Richard, and Ystad, Sølvi, editors, *Proceedings of the 16th International Symposium on Computer Music Multidisciplinary Research*. Tokyo, Japan, pages 450–461. https://doi.org/10.5281/zenodo.10113458.

Alejandro Enrique Planchart. 2017. On the nature of transmission and change in trope repertories. In *Embellishing the Liturgy*, Routledge, pages 167–201.

Leo Treitler. 2003. *With voice and pen: coming to know medieval song and how it was made*. OUP Oxford.

Günther Weiss. 1964. Zum problem der gruppierung südfranzösischer tropare. *Archiv fur Musikwissenschaft* pages 163–171.