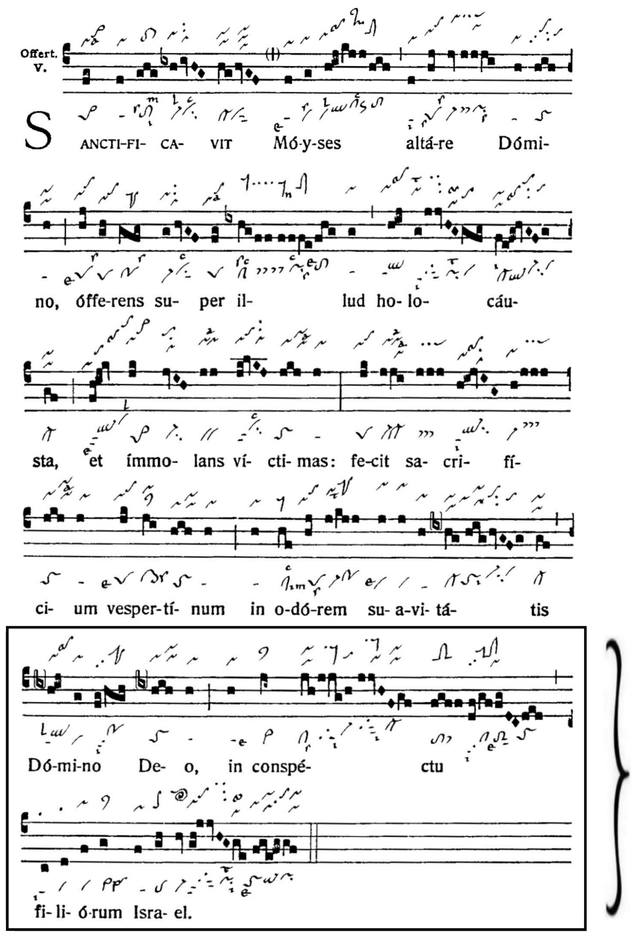
Connecting Medieval Chant Traditions: A Comparative Analysis

Peter van Kranenburg, Utrecht University & Geert Maessen, Gregoriana Amsterdam

ICCCM 2025, Aalborg University
Denmark, October 8-10, 2025

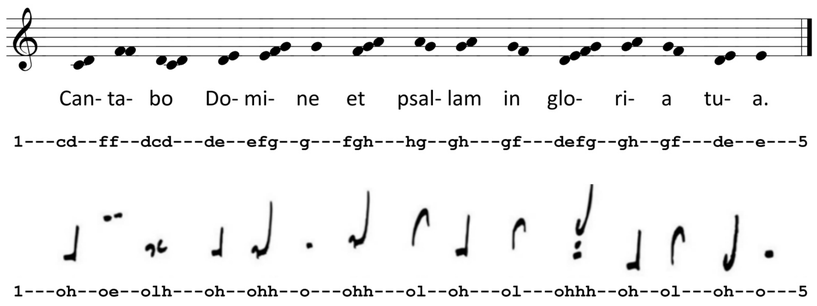
1. The first part/movement of GRE 64: *Sanctificavit Moyses*



2. Overview of our data: parts/movements in several traditions

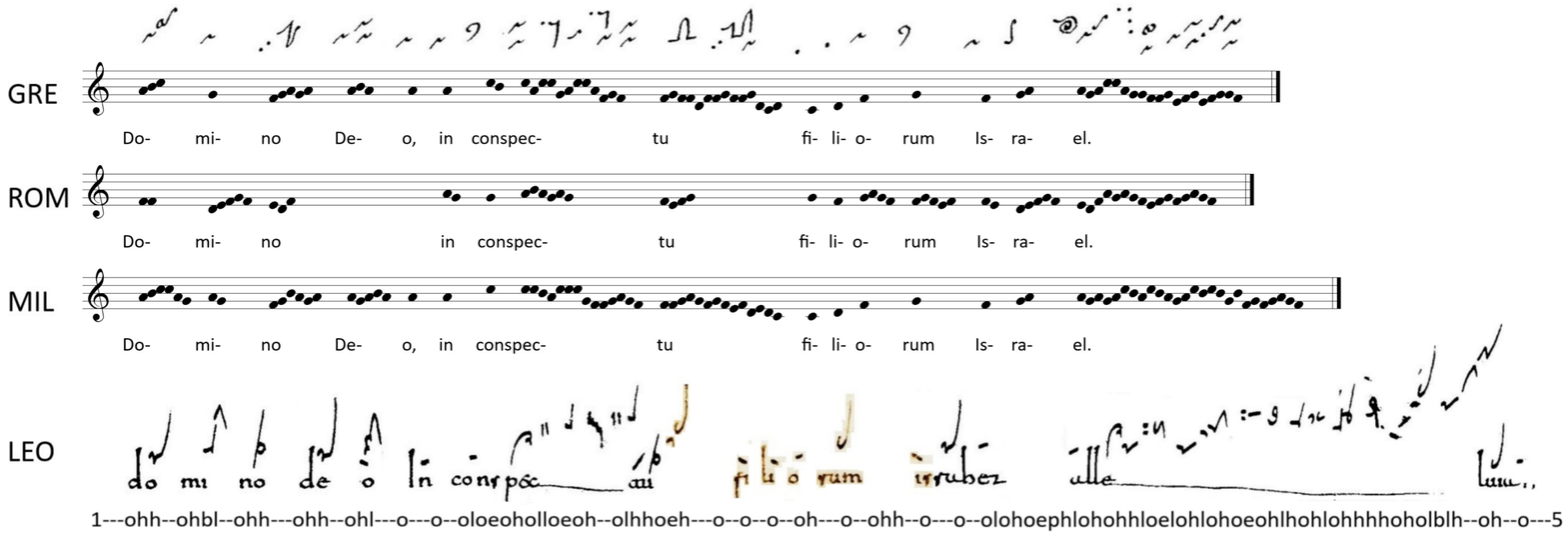
	chant tradition		century	chants	parts (p)	notes (n)	n/p
	GRE	Gregorian	X/XI	137	394	64123	163
		incl. GAL Gallician (?)	X/XI	40	103	18593	181
	ROM	Old Roman	XI/XII	94	285	48491	170
	MIL	Milanese	XI	104	147	26111	178
	BEN	Beneventan	X/XI	39	41	6272	153
Mozarabic	CIS	Cisneros Cantorales	XVI	71	139	17784	128
Mozarabic	HISA	Old Hispanic	XI	26	31	2343	76
Mozarabic	LEO	*León antiphoner*	*X*	*102*	*298*	*74967*	*252*

3. Encoding of pitch and *contour* notation



`o:` opening of a neume (indefinite),
`l:` lower than the previous note,
`b:` equal or higher,
`h:` higher than the previous note,
`e:` equal pitch
`p:` equal or lower

4. A passage in four traditions



5. Classification of chants by pitch and *contour*

For each tradition, we train n-gram language models. We assign an unknown chant to the tradition with the lowest perplexity. Left: a 5-gram model on the sequence of pitch intervals (in semitones). Right: a 2-gram model on the sequence of melisma lengths. Top: supervised evaluation (true class available for prediction). Bottom: forced misclassification (true class excluded; best among other classes chosen).

predicted \ true	GRE	ROM	MIL	BEN	CIS	HISA	parts
GRE	384	0	9	0	1	0	394
ROM	2	282	1	0	0	0	285
MIL	1	0	145	0	0	1	147
BEN	1	0	1	39	0	0	41
CIS	6	1	2	1	129	0	139
HISA	2	1	3	0	1	24	31
Precision	0.97	0.99	0.90	0.98	0.98	0.96	
Recall	0.97	0.99	0.99	0.95	0.93	0.77	
F1-Score	0.97	0.99	0.94	0.96	0.96	0.86	

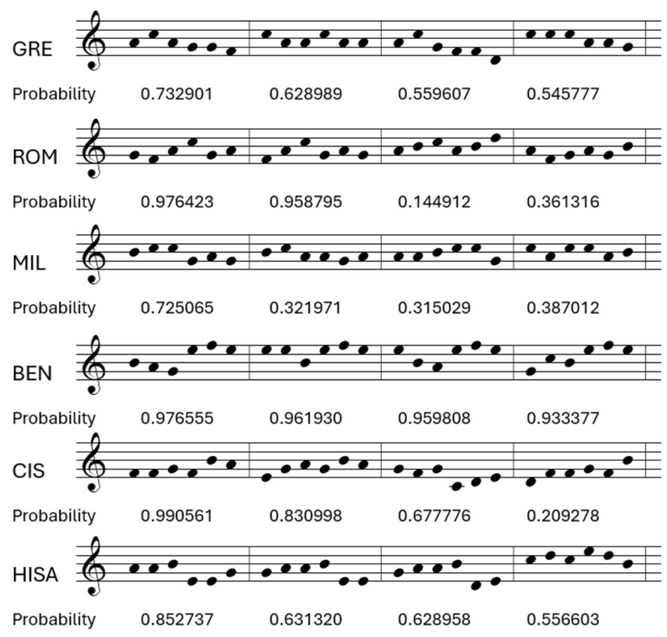
predicted \ true	GRE	ROM	MIL	BEN	CIS	HISA	parts
GRE	0	19	345	5	11	14	394
ROM	126	0	112	19	28	0	285
MIL	127	13	0	1	2	4	147
BEN	8	1	30	0	0	2	41
CIS	27	15	96	1	0	0	139
HISA	14	3	10	0	4	0	31

predicted \ true	GRE	ROM	MIL	BEN	CIS	HISA	LEO	parts
GRE	242	46	47	20	4	6	29	394
ROM	22	238	10	6	4	2	3	285
MIL	28	5	85	8	3	11	7	147
BEN	4	3	9	17	0	2	6	41
CIS	16	1	9	0	97	13	3	139
HISA	3	0	0	0	2	19	7	31
LEO	24	6	11	10	0	6	241	298
Precision	0.71	0.80	0.50	0.28	0.88	0.32	0.81	
Recall	0.61	0.84	0.58	0.41	0.70	0.61	0.81	
F1-Score	0.66	0.82	0.53	0.33	0.78	0.42	0.81	

predicted \ true	GRE	ROM	MIL	BEN	CIS	HISA	LEO	parts
GRE	0	93	153	39	11	8	90	394
ROM	160	0	35	49	5	4	32	285
MIL	67	12	0	35	6	14	13	147
BEN	8	7	15	0	0	3	8	41
CIS	30	6	60	0	0	39	4	139
HISA	3	0	1	0	13	0	14	31
LEO	195	27	25	27	0	24	0	298

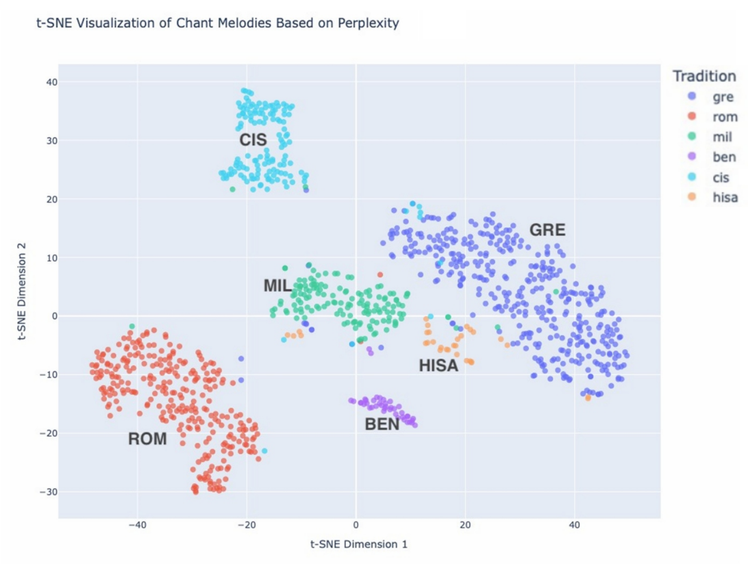
6. Distinctive six-note patterns in 6 traditions

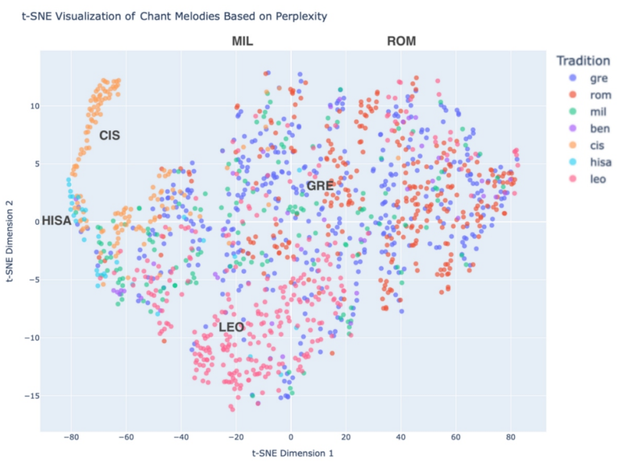
Top 5-grams for each chant tradition, selected for high distinctiveness (probability ≥10x higher than other traditions), significance (model probability >10^{-5}), and frequency (present in ≥5% of the tradition's melodies).

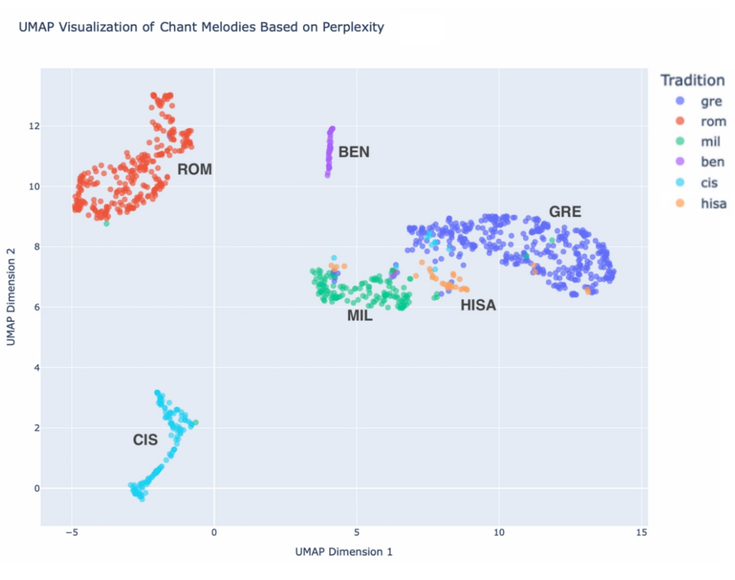


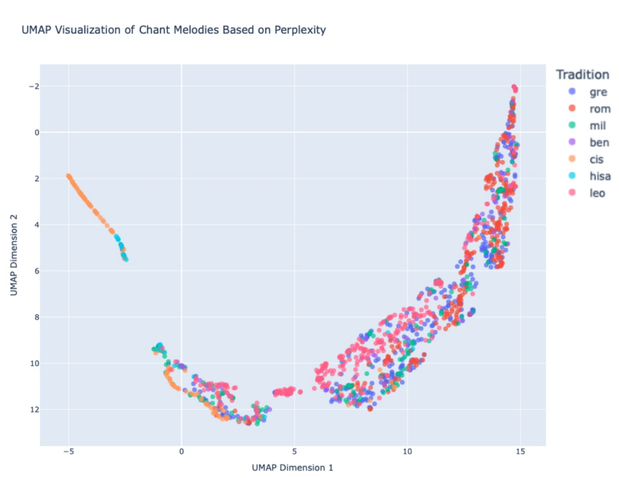
7. Dimension reduction for pitch and *contour*

For each melody in the various traditions, we compute for each tradition the perplexity using 5-gram models on pitch interval sequences, and 2-gram models on melisma-length sequences. We apply t-SNE and UMAP dimension reduction to visualize the perplexity space; left: scatter plot of reduced perplexities from 5-gram models, right: scatter plot from 2-gram models.





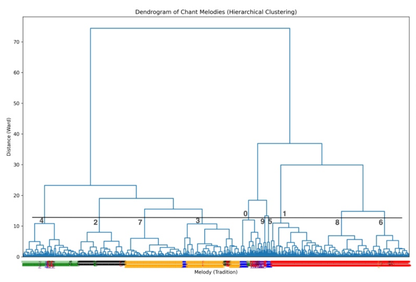




8. Clustering of sources, modes & Gallican chant

We perform hierarchical clustering in the perplexity-space.
Within the GRE-tradition, we discern three sub-clusters: clusters 1, 6, and 8.

tradition \ cluster	0	1	2	3	4	5	6	7	8	9
gre	10	140	1	6	6	0	98	0	127	6
rom	0	0	0	128	1	0	1	155	0	0
mil	9	0	5	1	131	0	1	0	0	0
ben	19	0	0	9	2	11	0	0	0	0
cis	2	0	120	10	3	0	4	0	0	0
hisa	21	0	0	0	6	0	0	0	0	4



We investigated whether these clusters align with different source manuscripts or with the modal structures of the melodies, but found no clear correspondence in either case. We also found no clearly distinguished cluster for the GAL subset within GRE.

source \ cluster	0	1	2	3	4	6	7	8	9
B	0	0	0	84	0	0	111	0	0
B34	5	81	0	2	5	51	0	75	1
B35	1	1	0	0	0	0	0	1	0
B39	0	9	0	0	0	1	0	2	0
L4951	0	1	0	0	0	0	0	2	0
M159	1	8	0	0	0	6	0	9	0
P1121	0	2	0	0	0	5	0	2	3
P776	1	38	1	4	1	32	0	36	2
P780	1	0	0	0	0	2	0	0	0
RoV52	1	0	0	0	0	1	0	0	0
V	0	0	0	44	1	1	44	0	0

mode \ cluster	0	1	2	3	4	6	7	8	9
1	1	24	0	16	0	11	24	21	0
2	2	18	1	16	5	18	26	18	1
3	2	23	0	19	0	9	15	12	0
4	1	38	0	39	0	10	14	13	3
5	1	4	0	6	0	9	16	11	0
6	1	5	0	4	0	15	17	10	0
7	0	4	0	5	0	4	1	2	0
8	2	24	0	29	2	23	42	40	2