Article

Generation of Melodies for the Lost Chant of the Mozarabic Rite

Darrell Conklin [1,2,*] and Geert Maessen [3]

[1] Department of Computer Science and Artificial Intelligence, University of the Basque Country UPV/EHU, 20018 San Sebastian, Spain
[2] IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Spain
[3] Gregoriana Amsterdam, Amsterdam, The Netherlands
[*] Correspondence: darrell.conklin@ehu.eus

Abstract: Prior to the establishment of the Roman rite with its Gregorian chant, in the Iberian Peninsula and Southern France the Mozarabic rite, with its own tradition of chant, was dominant from the sixth until the eleventh century. Few of these chants are preserved in pitch readable notation and thousands exist only in manuscripts using adiastematic neumes which specify only melodic contour relations and not exact intervals. Though their precise melodies appear to be forever lost it is possible to use computational machine learning and statistical sequence generation methods to produce plausible realizations. Pieces from the León antiphoner, dating from the early tenth century, were encoded into templates then instantiated by sampling from a statistical model trained on pitch-readable Gregorian chants. A concert of ten Mozarabic chant realizations was performed at a music festival in the Netherlands. This study shows that it is possible to construct realizations for incomplete ancient cultural remnants using only partial information compiled into templates, combined with statistical models learned from extant pieces to fill the templates.

Keywords: chant; Mozarabic rite; machine learning; statistical language models; music generation

1. Introduction

In medieval Europe several textually and musically related monophonic liturgical chant traditions existed. Most famous is the Franco-Roman chant of the Roman rite, better known as Gregorian chant. Most other rites and traditions were abolished at some point in favor of the Roman rite and its chant [1]. In 589 the Visigothic Kingdom of the Iberian peninsula was converted to Catholicism. In the early seventh century Iberian Catholicism developed into an independent rite of Christian worship which after the Muslim conquest of 711 became known as the Mozarabic rite. In 1080 this rite was officially abolished by the Council of Burgos and replaced by the Roman rite with its Gregorian chant. In 1085 Toledo, the centre of the Iberian church, was reconquered from Islam. Only six parishes of Toledo were allowed to continue the ancient rite. In the eleventh century pitch-readable music notation gradually came in use. Most chants of the Mozarabic rite, however, are only preserved in pitch-unreadable (adiastematic) neume notation [2]. The chants are preserved in about forty manuscripts and fragments dating from the early eighth until the thirteenth centuries. The most important manuscript is the León antiphoner (E-L 8, Catedral de León, MS 8), dating from the early tenth-century, containing over 3000 chants preserved in adiastematic neume notation.

Though the pitches of the melodies are unknown and probably lost forever, the neumes provide important information to assist in their realization: determination of a singable and plausible pitch sequence representing the neumes. The manuscripts, in neume notation with the syllables in the underlying text, provide two important pieces of information: the number of notes in each neume and

the melodic contour of the pitches internal to each neume. From note to note it is usually apparent if the melody goes up or down [3]. This contour information can be represented using six letters: h, a note higher than the previous note; l, lower; e, equal; b, higher or equal; p, lower or equal; o, a note with unclear and undefined relative height. Figure 1 shows a fragment of the *Canticum Zachariae* for the feast of St. John the Baptist. Shown at the top of the figure are two lines from the León antiphoner. Following that is the transcription of the neumes on the bottom line to contour letters. In the contour sequence syllables are separated by dashes and words by spaces. Finally the figure shows a passage of a performance score with a generated compatible melody (see Results).

Another important feature of chant is the presence of recurring *intra-opus* patterns within single chants [4] that would seem to represent the same melodic content in the lost chant, for example, the encircled neumes and bracketed contour sequence in Figure 1. There is a wide consensus among chant scholars that longer (i.e., 20 or more notes) intra-opus patterns do represent the same sequence of pitches [5]. Therefore in generated pieces, an *intra-opus* pattern should be instantiated by the same musical material. Repetition is ubiquitous in music and the generation of music containing repetitions is an important open topic in the area of music informatics, because it requires the solution of equality constraints between distant events in the music surface [6,7].
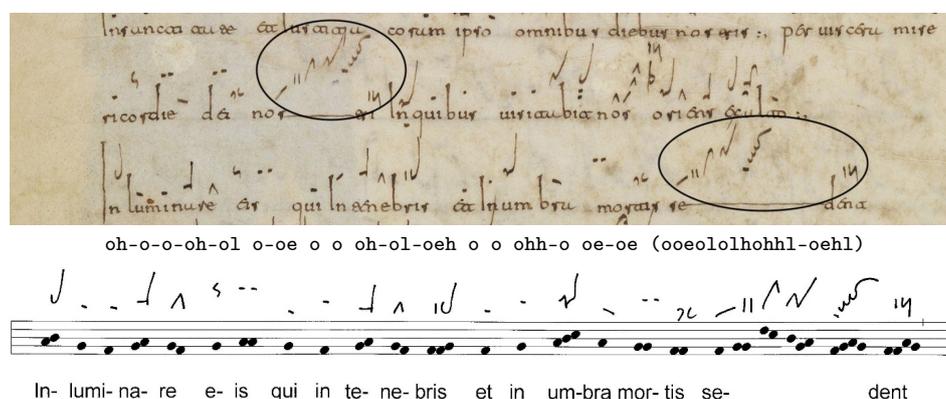

oh-o-o-oh-ol o-oe o o oh-ol-oeh o o ohh-o oe-oe (ooeololhohhl-oehl)

In- lumi- na- re   e- is   qui in te- ne- bris   et in um-bra mor- tis se-   dent

Figure 1. Two lines of the *Canticum Zachariae* for the feast of St. John the Baptist from the León antiphoner (E–L 8, 215r4–5). Encircled is an intra-opus repeating pattern. On the second line of the León image the beginning of the last verse; *Inluminare eis*. Below the León image is a representation in contour letters of this second line, and below that the corresponding passage of our generated performance score with text.

The core task of the adiastematic neume realization problem is to find pitches compatible with a specified template consisting of the melodic contour and intra-opus patterns. Since a vast number of melodies will be compatible with a given template, this is a highly under-constrained problem. Therefore a position must be taken on whether the task is viewed more as *restoration* or more as *generation*. Some scholars have shown melodic relations with other chant traditions for some specific Mozarabic chants [8]. In such cases chant realization may be approached as primarily a *restoration* task: using long fragments of concrete pitches found in a chant with known pitches and presumably with a historical relation. This is one motivation of the method of Maessen and van Kranenburg [9] for chant generation, which searches a corpus of preserved chants using contour descriptions of phrases from the template. If a closely matching database piece exists, it is used to overlay pitched fragments on the new chant. Remaining regions of the chant are constructed using less stringent matches. Finally manual editing of the borders between phrases will complete the melody. Even if a closely matching database chant exists, this method still requires expert intervention to fill in unmatched regions of the new chant [10]. The explicit use of long contour patterns drawn from a corpus has also been considered to be a general model for melody generation [11], where contour patterns specified by the composer, or selected from a predefined list, are instantiated by specific music segments drawn from a corpus.

This paper develops the alternative view of chant realization as primarily one of *generation*: making no a priori existence assumptions of closely related chants. Music generation approaches can broadly be grouped under rule-based (requiring specific rules and constraints to be encoded by the composer), and machine learning methods (learning rules and models from a training corpus) [12,13]. Most machine learning approaches to music generation use statistical models, originating from the earliest successful works with Markov models [14]. Most statistical models for music generation can be considered *context models*: generating the next event in a growing sequence based on the history of previously generated events. Context models encompass a wide range, including simple Markov models [15], n-gram and variable length Markov models [16], multiple viewpoint models [17], and deep learning models for music generation [18–21].

As mentioned above, a difficult problem for music generation methods is the precise control of intra-opus repetition, especially when using context models. There were some initial attempts to generate repetitive structures ab initio with context models [22]. An alternative powerful approach is to derive the repetition structure from known pieces, either automatically with intra-opus pattern discovery [23,24] or by compositional design. In this way the structure of a known piece is maintained in a newly generated piece. Things brings up the issue of how the structure is formally represented and instantiated: in this paper the method of Conklin [6], designed for generating chord sequences with complex repetition structures, was adapted to solve the chant realization problem.

An often overlooked aspect of statistical models for music generation is the sampling of solutions. A decision must be made whether a few solutions are found by optimization of the posterior probability (given specified information such as length and desired features of the generation), or whether a diversity of possible solutions is produced through random sampling from the posterior distribution of the statistical model [20]. For chant generation, given that there is no single "correct" realization of a template, it is important that diversity is attainable and that sampling methods are used to select from the vast space of possible sequences.

2. Method

This section describes the chant generation method: intra-opus patterns and their representation in a template, the statistical modelling and learning method, and the method for sampling new pieces from a statistical model.

2.1. Patterns and Templates

The chant realization problem can be modelled using templates that represent the desired attributes or *viewpoints* of the events at every position [6]. The set of viewpoints we use for chant are described in Table 1. Every viewpoint has a syntactic name, and a set of possible values it produces (its codomain). The pitch viewpoint describes the pitch of the event using a MIDI number. The position viewpoint is needed to index events in the sequence for contour computations. Following this are several contour viewpoints (see previous section for their semantics), each one a Boolean viewpoint (values t and f) specifying whether the indicated contour is satisfied. Finally a parameterized range viewpoint is used to specify, for each event, the lowest and highest pitches permissible.

Table 1. Viewpoints used in this study.

Viewpoint	Description	Codomain
pitch	set of 15 possible pitches	$\{57, 59, 60, \ldots, 81\}$
position	position of event in sequence	$\{1, 2, 3, \ldots\}$
h, l, e, b, p	contour viewpoints (see text)	Boolean
$\text{range}_{x,y}$	pitch in range $[x, y]$	Boolean

The composition of templates and their semantics is now described proceeding from the lowest level of *features* (attribute/value pairs) to entire templates. Example features are pitch:57, specifying

an event with the pitch 57, or $\text{range}_{57,72} : \text{t}$, specifying an event within the indicated range (see Figure 2c). A *feature set* represents a logical conjunction of features, for example $\{\text{e} : \text{t}, \text{range}_{60,67} : \text{t}\}$ representing an event with the same pitch as the previous note and with the specified range, or $\{\text{l} : \text{t}, \text{pitch} : 59\}$ which refers to an event with pitch 59, lower than the previous pitch. An event instantiates a feature set if it has all of the features in the set. A *template* is a sequence of feature sets, and is used to specify entire sequences with any desired properties at any location. A sequence instantiates a template if all successive events match the corresponding positions of the template from beginning to end.

Figure 2. (**a**) The first line of the responsory *Dominus ab utero* for the feast of St. John the Baptist in the León antiphoner (E-L 8, 214r2), with two intra-opus patterns; (**b**) a representation of the fragment in contour letters (with intra-opus patterns bracketed); (**c**) the partial encoding of the contour sequence and intra-opus patterns as a template. $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are variables used to specify equal pitches; (**d**) a random instantiation of the template: the two different colors indicate the beginning event for the two intra-opus patterns; (**e**) a sample taken from a statistical model, ignoring any template constraints; (**f**) the corresponding passage of our generated performance score.

To specify the sharing of values among different events, necessary for specifying intra-opus patterns and long range dependencies, the notion of a feature is extended to include *variables*. For example, the feature $\text{pitch} : \mathcal{V}$ can be used to specify an event with some variable pitch $\mathcal{V}$, and the variable can occur elsewhere in the sequence. In Figure 2c, variables $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are used for specifying equal pitches. The variable $\mathcal{A}$ occurs at the first event of each occurrence of the first intra-opus pattern. Please note that the first occurrence of $\mathcal{A}$ also has a defined pitch, which by implication fixes the second occurrence to the same pitch.

To give the semantics of templates with variables, the notion of a *substitution* is required. A substitution is a function from variables occurring in a template to elements of the codomains of viewpoints appearing within the template. Thus a substitution applied to a template produces a sequence with all variables instantiated by concrete pitches. Every different variable substitution will produce a different event sequence. If a sequence e instantiates a template Φ under some variable substitution, this is notated $\Phi(e) = \text{t}$, using the syntax $\Phi(e)$ so that Φ can be interpreted as both a template and a Boolean function of event sequences.

2.2. Statistical Model

The core task of chant realization can be viewed as instantiating a given template Φ with compatible sequences. Compatibility with a template, however, is a necessary but not a sufficient condition for good generated music. Selecting arbitrary sequences that instantiate a template is highly

unlikely to generate good musical material. To see this, consider the template shown in Figure 2a–c. The score fragment in Figure 2d is a random sequence instantiating the template. This is a poor sequence, hardly singable and containing too many large leaps. Its information content (IC), measured in terms of average negative log (base 2) probability per event, is high (5.8 bits/event) which indicates a low probability sequence according to the statistical model. It wanders excessively between low and high parts of the range. On the other hand, the third fragment in Figure 2f—a passage of a melody produced for a generated performance score (see Results)—was generated by taking 1000 compatible samples (see following subsection) using a statistical model trained on a corpus of chant melodies. Its information content is low (1.95 bits/event). It can be seen that the melodic line is smooth while still respecting the template. These two fragments illustrate the importance of complementing a template with a statistical model.

The score fragment in Figure 2e is generated by sampling a low IC sequence (1.8 bits/event) using the same statistical model though without any intra-opus patterns or contours specified in the template, using only the pitch ranges and first defined pitch. This sequence, while having high probability according to the model, is also poor as it hovers around just a few pitches. This illustrates the importance of complementing a statistical model with a template.

More precisely, a statistical model assigns a probability $P(e_1, \ldots, e_n)$ to a sequence of events $e = e_1, \ldots, e_n$. A *context model* factors this joint probability into a sequence of probabilities for each event, each conditioned on k events of history:

$$P(e) = \prod_{i=1}^{n} P(e_i \mid e_1, \ldots, e_{i-1}) \approx \prod_{i=1}^{n} P(e_i \mid e_{i-k+1}, \ldots, e_{i-1}) = \prod_{i=1}^{n} P(e_i \mid h_i) \tag{1}$$

where h_i stands for the history (context) of the event e_i. For chant generation here an efficient yet powerful context model called PPM [25] is employed. These types of models were highly effective in the past for music modeling [17,26] as they have the ability to capture important local dependencies, and to reach back further in time by interpolating contexts of variable lengths. They are variable-order n-gram models which are learned by compiling an indexed dictionary. For prediction after learning, they interpolate the probabilities of different context lengths (up to a maximum length k) together to produce a final probability for an event. Here the simple backoff variant [27] is used (see Appendix A.5): progressively backing off to lower order contexts, at every stage multiplying in an escape probability computed from the history, until a match can be found in the dictionary. Many other variants of PPM are possible [26], as are other types of statistical context models. The aim here for chant generation is not to search for the optimal statistical model but rather to rely on the use of specific templates that will moderate even an underfit model.

2.3. Sampling Compatible Instances of Templates

Context models, though practical and efficient for prediction tasks, cannot capture nonlocal repetition in the music surface and therefore alone be expected to generate good musical structures. They can however be combined with designed templates that specify the necessary and desired structure. Let E be a random variable ranging over sequences, and Φ be the Boolean random variable indicating whether a sequence instantiates the given template Φ. Using Bayes' rule, the likelihood of a sequence e, given that the template Φ is instantiated, is provided by:

$$P(E\!=\!e \mid \Phi\!=\!\mathsf{t}) = \frac{P(\Phi\!=\!\mathsf{t} \mid E\!=\!e) \times P(E\!=\!e)}{P(\Phi\!=\!\mathsf{t})} \propto P(\Phi\!=\!\mathsf{t} \mid E\!=\!e) \times P(E\!=\!e) \tag{2}$$

with the proportionality holding because the denominator is a normalizing constant, representing the proportion of all sequences instantiating the template, and it depends only on the template.

In Equation (2) the marginal probability $P(E = e)$ is defined by Equation (1) and the likelihood of a template given an event sequence is given by a Bernoulli distribution:

$$P(\Phi = \text{t} \mid E = e) = \begin{cases} 1 & \Phi(e) = \text{t} \\ 0 & \Phi(e) = \text{f} \end{cases} \tag{3}$$

which states that templates are either instantiated or not (i.e., there is no gradation of instantiation).

Generating single solutions from a model, given a template Φ, is performed according to Equation (2) by sampling from the distribution $E \mid \Phi = \text{t}$. This reduces to sampling from the right hand side of Equation (2). Algorithmically, for the type of templates used here for chant generation, the problem can be solved with random walk [6] combined with constraint satisfaction methods. Sequences are generated left-to-right while maintaining a partial variable substitution μ. The substitution μ is initially empty and is updated every time a variable is instantiated. The substitution μ and the feature set at a template position i determines dynamically the set of permissible events $dom_\mu(\Phi_i)$. To generate a sequence $e = e_1, \ldots, e_n$ using random walk, we proceed left-to-right, sampling events $e_i \in dom_\mu(\Phi_i)$ with the probability $P(e_i \mid h_i)$, appropriately normalized to the probability mass of $dom_\mu(\Phi_i)$. This procedure can be performed without backtracking, provided that the underlying statistical model is *non-exclusive*: assigning a probability, however tiny, to every possible event at every position.

A known issue with the random walk method is that while it produces a diversity of valid solutions, it does not sample exactly from the distribution $E \mid \Phi = \text{t}$ for complex templates such as the ones used for chant, which can express equality relations [6,7]: the expected number of samples of a pattern instance e in n iterations does not converge to $n \times P(e \mid \Phi = \text{t})$. This happens because random walk performs no lookahead, and peaks of high IC are encountered during the left-to-right sampling procedure. One way to address this issue, potentially more accurate though needing higher computational resources, is by using approximate Monte Carlo methods for sampling, for example iterative random walk [6] to generate a large number of solutions by restarting the random walk many times. Sequences can be subsequently selected from the distribution of all distinct sequences sampled and retained during the iterations. This is the method employed here for chant generation.

3. Results

This section describes the application of the chant generation method to produce an entire concert of generated pieces. First the training corpus is described, followed by the creation of templates for several Mozarabic chants. The properties of the core statistical model are outlined, followed by a description and audience evaluation of the concert pieces.

3.1. Corpus

A corpus of 137 Gregorian offertories (GRE) in pitch-readable notation was used to train statistical models on absolute pitches occurring in the corpus (see Appendix A.1). The corpus has approximately 65,000 notes; Table 2 provides (top) some descriptive statistics of the corpus. Of five different chant traditions Gregorian chant appears to be the most similar to the lost chant of the León antiphoner [28]. Since the manuscript sources do not provide information about rhythm and metre, this information is not included in the corpus encoding. Though the data is purely symbolic, for the convenience of contour computations MIDI numbers can be used as there are no enharmonics. Furthermore since the accidentals on the notes B♭ and E♭ are not consistently encoded in the corpus, they are ignored (the notes are considered to be the notes B and E respectively) during training and generation.

Table 2. Description of the GRE corpus (top) and the 22 templates (bottom).

number of chants in GRE	137
mean chant length	473 notes
mean number of words/syllables/neumes	56/123/318
number of templates	22
mean template length	789 notes
mean number of defined pitches	14
mean number of words/syllables/neumes	107/226/464
mean coverage by intra-opus patterns	52%
mean fraction of events with no specified contours	34%

3.2. Template Creation

Templates were compiled for 22 Mozarabic chants from the León antiphoner (see Appendix A.2). Reasons for the choice of these 22 pieces include: the complexity of the chants; a representative selection from the León antiphonary; and possibilities for different thematic units for performance. The choice included 12 sacrificia, 8 responsories, a sono and a Benedictus, many for Easter time, and ten chants for the feast of St. John the Baptist. As described above, these templates encode neume contours, intra-opus patterns, ranges, and some defined pitches. Templates for each piece were based on enhanced digital images of the León antiphonary by manually representing the neumatic notation of León in contour letters. This was based on the findings of Rojo and Prado [3] and further on the work of González-Barrionuevo [29] who described the meaning of the neumes and partially their interpretation.

For the correct transcription of words and syllables we made use of the text edition by Brou and Vives [30]. We carefully looked for intra-opus patterns—sequences of neumes that are repeated—and manually marked these using brackets (see Figures 1 and 2). First and last pitches, as well as ranges, were arbitrarily chosen based on Gregorian chants on the same places of the liturgical calendar [31,32]. Randel [33] associates the verses of nearly 600 responsories with one of 7 psalm tones: A, B, C, D, E, F, and G. Since the pitches of two tone-B verses are known (of the responsories *Ecce ego viam* and *Dies mei transierunt*), most pitches of our four tone-B verses (of *Haec dicit Dominus iustitia*, *Zaccarias*, *Unde mici* and *Me oportet*) were defined by these. Because the actual verse texts determine the neumatic structure of these verses, different tone-B verses can differ considerably, although they are closely related. Therefore it was not possible to define all pitches of our tone-B responsory verses.

Most of the 22 pieces have *repetendae*, longer parts of the chant that should be repeated after a verse. In the manuscript the repetendum is only copied the first time and the subsequent occurrences are simply indicated with the first word of the text. Our encoding, however, copied them always completely, thus creating longer intra-opus patterns. Responsories nearly always have the general form I-R-V-R, and sacrificia I-R-II-R-III-R where I is the initium, R the repetendum, V the verse, and II and III the second and third parts. To conform to the melodic behaviour of the related genre of the offertory in Gregorian chant, we assigned a different range to the final (third) part of sacrificia (SCR) compared to the rest. Since the offertory is the only genre with this feature, the range for all other genres is the same throughout the piece.

Table 2 (bottom) provides some descriptive statistics of the template set. These templates provide very challenging generation tasks, with long sequences, many defined pitches, and over one-half of all notes covered by intra-opus patterns that must be respected in generated pieces. One-third of the events on average are under no specified contour constraint, thus necessitating a good statistical model to compensate and choose good melodic material for these positions.

3.3. Statistical Model

For prediction several PPM(k) models were trained on the GRE corpus. Figure 3 (left) shows the average information content, cross-validated by leave-one-out analysis, for four models of different

orders: 0 (unigram), 1 (bigram), 2 (trigram), and 4 (pentagram). It can be seen that the models, while decreasing the information content as desired, progressively overfit the data as more surprising (high information content) events are encountered. The pentagram model, for example, appears to overfit the corpus more than models of lower order, seen by the longer tail of high IC events. The trigram model is considered a good balance between bias and variance and is used as the base statistical model for chant generation.
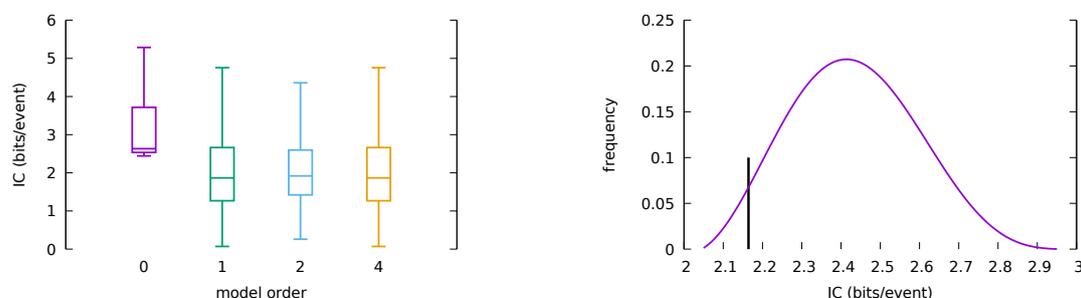

Figure 3. (**Left**): leave-one-out per event information content of the GRE corpus, under different orders of PPM(k) model. (**Right**): the distribution of sequences generated by 10,000 iterations of random walk, for a template of 840 events and a PPM(2) model of the GRE corpus. The vertical black line marks the mean IC to the training corpus.

Following model training, a sequence of pitches can be generated based on the probabilities derived from the data set by performing statistical sampling and settling on sequences at the high end of probability space. Figure 3 (right), for the chant *Dominus ab utero*—the full 840 note piece for the fragment in Figure 2—sequence probabilities produced by 10,000 iterations of random walk. In those samples, 9673 unique sequences were generated, showing that iterative random walk produces a high diversity of sequences. It can be seen that the information content (here divided by the number of events) follows an extreme value distribution, with a longer tail of low probability pieces. The black vertical line indicates the mean IC to the training corpus, showing that it lies in the low IC (high probability) tail of the sampled distribution, and well under the mean of the sampling distribution. This is due to two factors: random walk is a biased sampling procedure in the presence of complex patterns; and requiring template instantiation can skew the distribution towards lower probability sequences.

3.4. Concert of Generated Chants

The method was employed to generate an entire suite of chants that were performed at the Nederlands Gregoriaans Festival, 's-Hertogenbosch, on 14–16 June 2019. From the 22 encoded pieces, a smaller set of 10 was chosen (Table 3): one, *Dum complerentur*, for Pentecost (9 June) and nine for St. John (24 June). For each of these templates 1000 iterations of random walk were performed to produce high probability solutions, of which simply the highest was chosen for the concert. Only two minor edits were made by hand in a single chant, *Benedictus Dominus*, to break undesirable sixth and fifth leaps. See Appendix A.3 for links to the entire scores of the concert pieces, and Appendix A.4 for links to audio recordings.

For the concert the ensemble *Gregoriana Amsterdam* consisted of four professional singers and the director who is an expert in chant performance. The order of the pieces was changed from the León antiphoner in order to get a running story. Therefore also three short liturgical lectures were included in the concert. For reasons of performance variety not all chants were treated the same. The repetendae were always sung by all five (tutti), but the other parts were sung in different combinations. The *Benedictus Dominus*, which consists of ten verses without repetendae, was sung alternatim; the odd verses by a solist and the even tutti. The rhythmic interpretation of the León

neumes was inspired by the semiological interpretation of Gregorian chant [34] as sketched by González-Barrionuevo [29].

Table 3. Chants performed at the concert, including genre (SNO: sono; RS: responsory; VAR: various; SCR: sacrificium), performance time, name of template, place in E-L 8, and for the generated performance score (see Results), IC (bits/event), audience and singer ratings

Genre	Time	Incipit	E-L 8	IC	Audience (n = 34)		Singers (n = 5)	
					Mean	Stdev	Mean	Stdev
SNO	07:03	*Haec dicit Dominus priusquam*	211v10	2.03	7.5	1.8	6.2	0.8
RS	04:38	*Zaccarias sacerdos*	212v11	2.03	7.6	1.7	6.6	0.5
RS	02:05	*Unde mici adfuit ut veniret*	213v10	2.08	8.0	1.6	5.8	1.8
RS	02:31	*Fuit homo missus a Deo*	213v01	1.88	8.2	1.3	7.2	0.8
RS	03:47	*Dominus ab utero formabit me*	214r02	1.95	7.7	1.6	7.8	1.1
RS	02:14	*Spiritus Domini super me*	213r08	1.99	8.0	1.6	8.0	0.8
RS	02:53	*Misit me Dominus sanare*	212v02	1.99	8.2	1.4	7.5	1.4
RS	02:16	*Me oportet minui*	214r12	1.96	8.9	1.4	6.4	0.9
VAR	07:06	*Benedictus Dominus Deus Israel*	214v12	2.08	9.4	1.0	5.9	2.5
SCR	07:22	*Dum complerentur dies*	210r14	2.12	8.7	1.4	6.2	0.8

3.5. Singer and Audience Evaluation

Before the concert a questionnaire was handed to the audience to obtain feedback on the concert pieces. Of approximately 50 to 60 attendees, a total of 34 people completed the form. All respondents had a specific interest in chant. Many are musically trained, as singers, directors or music teachers, some even as researchers. In the form were five choices to evaluate a piece (in Dutch, here with rough English equivalents); niks (poor), zwak (weak), neutraal (neutral), aardig (nice), and prachtig (beautiful). These categorical scores were converted to numeric scores 2, 4, 6, 8, and 10. This conversion allows some rough comparison with the singer evaluation which was a numeric score in the range 0–10 (below). Mean and standard deviation of audience responses are presented in Table 3. It can be seen that the means fall mostly in the range "nice" to "beautiful", with some pieces (notably *Benedictus Dominus*) receiving high scores and overall low deviation.

Apart from their evaluation score, 22 of the respondents used the back of the form to write their observations. Twelve of these provided information about specific chants. Four people observed that it was difficult to discriminate between the pieces and the performance. Four others in fact only made observations about the performance, despite the questionnaire asking listeners to focus on the melodies. Five people found all chants similar. Since this is often observed in chant concerts this, also, does not tell us anything about the melodies themselves. As with other chant traditions, however, it could also be seen as an indication that these people simply were not able to hear the nuances of the melodies. Five people also stated that they preferred fewer singers instead of the full choir. Again, this gives no information about the melodies. We can see these facts reflected in Table 3: *Me oportet minui* and *Benedictus Dominus* were largely sung by a soloist, and *Dum complerentur* almost entirely by two singers. These three chants received the highest audience scores.

The five singers were prepared for the concert in four rehearsals. All five are professionals with great experience in early music. However, each of them has a specific expertise, be it as composer, choir director, choir singer, liturgical singer or researcher. After the third rehearsal, when all ten chants were well known, the singers were asked to evaluate the chants, with a score between 0 (very poor) and 10 (very good). Mean and standard deviation of their evaluation is included in Table 3. The rating of the singers had more variance and the highest scores were given to different pieces than by the audience. The difference between the rating of the audience and the singers can be understood in several ways. First, the singers did not need to distinguish between the performance and the melodies. Therefore they were able to focus on the melodies themselves. Secondly, they definitely had their specific biases based on their specific expertise as composers, directors, singers and church musicians, which can explain

the differences in standard deviations, especially for the *Benedictus Dominus*. Third, the transformation from the 5-category qualitative evaluation of the audience evaluations into a numeric score in the range 0–10, though permitting a rough comparison with the singer scores, has naturally introduced some incongruency. Finally, the difference between the audience mean and that of the singers can be understood simply from the setting. The audience was enjoying a concert, while the singers were at work.

4. Discussion

This paper described a new method for chant generation which explicitly conserves the structure present in defined templates. Templates were carefully designed using musicological considerations and a statistical model learned from presumably related musical material was used to instantiate the templates. The method was used to generate an entire concert suite of chants which was performed at a music festival in the Netherlands.

The research has opened up two interesting issues, both arising late in the process while a concert suite was in the final stages of generation. The first issue concerns high information peaks which happen when the start of an intra-opus pattern or a defined pitch is encountered during a left-to-right random walk. In these cases the sequence might have to return to an instantiated event with an unnatural leap and low probability. This issue arises with random walk on complex templates, and an exact solution is possible only for the simplest types of statistical models and templates such as first-order Markov models with unary constraints on positions [15]. The information peak issue can produce low probability sequences because in the presence of complex templates, it is difficult or intractable to sample sequences with the same expected frequency as defined by their probability according to the statistical model. Several inexact methods were proposed, such as Gibbs sampling [19], bi-directional LSTM models [21], and iterative random walk as applied in the present paper [6].

A second unanticipated issue that arose is that melodies had a tendency to sit in the upper range for too long. This was observed by all the singers, although only mentioned in 3 of the 22 written audience comments. The phenomenon arises due to the presence of many undefined contours in templates, combined with the very slight preference in the statistical model for upwards contours. This can be corrected by limiting the number of undefined contours in templates, for example, by replacing them by either a concrete contour to the previous neume, or contour relation to the first note of the previous neume. Indeed inter-neume contour relations can sometimes be inferred from the manuscripts [4]. Another solution to this problem could be the generation of entire neumes rather than single notes. Here, however, it is possible that data sparsity problems would arise for model training.

A fascinating point opened up by our research is the role of overfitting in statistical models. Usually overfitting is viewed entirely negatively as the inability of a model to generalize past the known data. However in the chant realization problem there are cases where overfitting is desired. If *restoration* is desired and there is a closely related chant in the corpus, an overfit model should be able to retrieve long fragments from that chant whereas a model trained for *generation* will tend to mainly generate novel material. It is hypothesized that statistical methods can handle both sides of the spectrum, trained to fit to any degree the training corpus, including memorizing long fragments from the corpus.

Automated pattern discovery algorithms [35] might be used to find intra-opus patterns in the template contour sequence, thus automating the laborious step of hand annotation of a template for intra-opus patterns. Interesting patterns could be determined by statistical significance measures. To create a large collection of realizations for many templates the application of automated pattern discovery seems even necessary. An important extension of this work will be to consider *inter-opus* patterns, i.e., patterns appearing across different pieces within a corpus of template pieces. If the generation problem is viewed as one of generating a suite of pieces, it is desirable that the generated pieces have some inter-opus coherence. If inter-opus patterns are detected in different pieces in the manuscript they should also be instantiated with similar musical material in generated pieces.

5. Conclusions

This paper presented an approach to the realization of plausible melodies for lost chants of the Mozarabic rite. Templates are created from manuscripts, and contain information related to melodic contour and intra-opus patterns. A general statistical model, trained from a corpus of pitched chants, is used to produce high probability instantiation of templates via an iterative random walk sampling scheme. It is hypothesized that this general approach could be used for musicological studies and generation of other corpora, and could even be extended beyond music to the realization of lost linguistic and phonetic texts.

Author Contributions: D.C. designed and implemented the generation method. G.M. created the corpus and the templates. Both authors designed the experiments, analyzed the results, and wrote and reviewed the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

GRE   Gregorian corpus
PPM   Prediction by Partial Match
LSTM   Long Short-Term Memory
IC   Information Content

Appendix A. Supplementary Information

Appendix A.1. GRE Corpus

The GRE corpus consists of 137 pieces: all offertories found in three eleventh century manuscripts (F-MOf: H0159; F-Pn: Ms Lat 00776 and I-BV: Ms 34). The manuscripts are available through the Medieval Music Manuscripts Online Database (http://musmed.eu/sources).

Appendix A.2. 22 Templates

The 22 templates can be found at http://www.gregoriana.nl/templates-of-22-chants.txt.

Appendix A.3. Scores for Concert Pieces

The scores for the concert pieces can be found at http://www.gregoriana.nl/scores-for-16-june-2019.pdf.

Appendix A.4. Links to Recordings

The concert with 10 pieces was performed on 16 June 2019, recorded live by Concertzender and broadcasted on Friday 9 August. The recording is available at: https://www.concertzender.nl/programma/concertzender_live_518833/. Some of the pieces were also uploaded as videos to the Internet, where the live recording is complemented with the synchronous playing of the manuscript images: http://www.gregoriana.nl/videos.htm.

Appendix A.5. PPM Backoff Model

The backoff variant of the PPM(k) model is described. We assume here that there are no out-of-vocabulary events, i.e., for all events their count $c(e)$ in the corpus is greater than 0. Recall

from Equation (1) that $P(e) \approx \prod_1^n P(e_i \mid h_i)$. The backoff PPM($k$) model computes $P(e \mid h)$ using the following recurrence (starting h from the longest available context for e, of length no more than k):

$$P(e \mid h) = \begin{cases} P(e \mid s(h)) & c(h) = 0 & \text{(history never seen)} \\ P(e \mid s(h)) \times \gamma(h) & c(h) > 0 \text{ and } c(he) = 0 & \text{(event never seen with this history)} \\ c(he)/c(h) \times (1 - \gamma(h)) & c(h) > 0 \text{ and } c(he) > 0 & \text{(event seen with this history)} \end{cases}$$

where $s()$ returns the longest proper suffix, i.e., $s(e_1, \ldots, e_n) = e_2, \ldots, e_n$ and where $\gamma(h)$ is the backoff (escape) of h: the probability mass assigned to events not seen before in the context of h. With Method C discounting [36] this is $\gamma(h) = \frac{u(h)}{u(h)+c(h)}$ where $u(h)$ is the number of unique letters following h.

References

1. Hiley, D. *Gregorian Chant*; Cambridge University Press: Cambridge, UK, 2009.
2. Randel, D.M.; Nadeau, N. Mozarabic Chant. 2001. Available online: http://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000019269 (accessed on 15 April 2019).
3. Rojo, C.; Prado, G. *El Canto Mozárabe, Estudio Histórico-critico de su Antigüedad y Estado Actual*; Diputación Provincial de Barcelona: Barcelona, Spain, 1929.
4. Maessen, G. Aspects of melody generation for the lost chant of the Mozarabic rite. In Proceedings of the 9th International Workshop on Folk Music Analysis (FMA 2019), Birmingham, UK, 2–4 July 2018; pp. 23–24.
5. Hornby, E.C.; Maloy, R. Toward a Methodology for Analyzing the Old Hispanic Responsories. In *Cantus Planus Study Group of the International Musicological Society*; Österreichische Akademie Der Wissenschaften: Vienna, Austria, 2012; pp. 242–249.
6. Conklin, D. Chord sequence generation with semiotic patterns. *J. Math. Music.* **2016**, *10*, 92–106. [CrossRef]
7. Rivaud, S.; Pachet, F.; Roy, P. Sampling Markov Models under Binary Equality Constraints is Hard. In *Journées Francophones sur les Réseaux Bayésiens et les Modéles Graphiques Probabilistes*; csl.sony.fr: Clermont-Ferrand, France, June 2016.
8. Levy, K. *Gregorian Chant and the Carolingians*; Princeton University Press: Princeton, NJ, USA, 1998.
9. Maessen, G.; van Kranenburg, P. A Semi-Automatic Method to Produce Singable Melodies for the Lost Chant of the Mozarabic Rite. In Proceedings of the 7th International Workshop on Folk Music Analysis (FMA 2017), Malaga, Spain, 14–16 June 2017; pp. 60–65.
10. Maessen, G.; Conklin, D. Two methods to compute melodies for the lost chant of the Mozarabic rite. In Proceedings of the 8th International Workshop on Folk Music Analysis (FMA 2018), Thessaloniki, Greece, 26–29 June 2018; pp. 31–34.
11. Roig, C.; Tardón, L.J.; Barbancho, I.; Barbancho, A.M. Automatic melody composition based on a probabilistic model of music style and harmonic rules. *Knowl.-Based Syst.* **2014**, *71*, 419–434. [CrossRef]
12. Conklin, D. Music generation from statistical models. In Proceedings of the AISB Symposium on Artificial Intelligence and Creativity in the Arts and Sciences, Brighton, UK, 7–11 April 2003; pp. 30–35.
13. Fernandez, J.D.; Vico, F.J. AI Methods in Algorithmic Composition: A Comprehensive Survey. *J. Artif. Intell. Res.* **2013**, *48*, 513–582. [CrossRef]
14. Brooks, F.P.; Hopkins, A.L., Jr.; Neumann, P.G.; Wright, W.V. An experiment in musical composition. *IRE Trans. Electron. Comput.* **1956**, *EC–5*, 175–182. [CrossRef]
15. Pachet, F.; Roy, P.; Barbieri, G. Finite-length Markov processes with constraints. In Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011), Barcelona, Spain, 16–22 July 2011; pp. 635–642.
16. Dubnov, S.; Assayag, G.; Lartillot, O.; Bejerano, G. Using Machine-Learning Methods for Musical Style Modeling. *IEEE Comput.* **2003**, *36*, 73–80. [CrossRef]
17. Conklin, D.; Witten, I. Multiple viewpoint systems for music prediction. *J. New Music. Res.* **1995**, *24*, 51–73. [CrossRef]
18. Sturm, B.L.; Santos, J.F.; Ben-Tal, O.; Korshunova, I. Music transcription modelling and composition using deep learning. *arXiv* **2016**, arXiv:1604.08723.

19. Huang, C.A.; Cooijmans, T.; Roberts, A.; Courville, A.; Eck, D. Counterpoint by convolution. In Proceedings of the 18th International Society for Music Information Retrieval Conference, Suzhou, China, 23–27 October 2017; pp. 211–218.

20. Walder, C.; Kim, D. Computer assisted composition with Recurrent Neural Networks. *JMLR: Workshop Conf. Proc.* **2017**, *80*, 1–16.

21. Hadjeres, G.; Nielsen, F. Anticipation-RNN: Enforcing unary constraints in sequence generation, with application to interactive music generation. *Neural Comput. Appl.* **2018**. [CrossRef]

22. Medeot, G.; Cherla, S.; Kosta, K.; McVicar, M.; Abdallah, S.; Selvi, M.; Newton-Rex, E.; Webster, K. StructureNet: Inducing Structure in Generated Melodies. In Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR 2018), Paris, France, 23–27 September 2018; pp. 725–731.

23. Cope, D. *Virtual Music: Computer Synthesis of Musical Style*; The MIT Press: Cambridge, MA, USA, 2001.

24. Collins, T.; Laney, R.; Willis, A.; Garthwaite, P.H. Developing and evaluating computational models of musical style. *Artif. Intell. Eng. Des. Anal. Manuf.* **2016**, *30*, 16–43. [CrossRef]

25. Cleary, J.G.; Witten, I.H. Data compression using Adaptive coding and Partial String Matching. *IEEE Trans. Commun.* **1984**, *32*, 396–402. [CrossRef]

26. Pearce, M.T.; Wiggins, G.A. Improved methods for statistical modelling of monophonic music. *J. New Music. Res.* **2004**, *33*, 367–385. [CrossRef]

27. Chen, S.F.; Goodman, J. An empirical study of smoothing techniques for language modeling. *Comput. Speech Lang.* **1999**, *13*, 359–393. [CrossRef]

28. Maessen, G.; van Kranenburg, P. A Non-Melodic Characteristic to Compare the Music of Medieval Chant Traditions. In Proceedings of the 8th International Workshop on Folk Music Analysis, Thessaloniki, Greece, 26–29 June 2018; pp. 78–79.

29. González-Barrionuevo, H. The Simple Neumes of the León Antiphonary. In *Calculemus et Cantemus, Towards a Reconstruction of Mozarabic Chant*; Maessen, G., Ed.; Gregoriana: Amsterdam, The Netherlands, 2015; pp. 31–52.

30. Brou, L.; Vives, J. (Eds.) *Antifonario visigótico Mozárabe de la Catedral de León (Monumenta Hispaniae Sacra Serie Litúrgica, Vol. V,1)*; Consejo Superior de Investigaciones Científicas: Madrid, Spain, 1959.

31. Billecocq, M.C.; Fischer, R. (Eds.) *Graduale Triplex*; Abbaye Saint-Pierre: Solesmes, France, 1979.

32. Ott, K.; Fischer, R. (Eds.) *Offertoriale Triplex cum Versiculis*; Abbaye Saint-Pierre Solesmes: Solesmes, France, 1985.

33. Randel, D. *The Responsorial Psalm Tones for the Mozarabic Office*; Princeton University Press: Princeton, NJ, USA, 1969.

34. Cardine, E. *Semiologia Gregoriana*; Pontificium Institutum Musicae Sacrae: Rome, Italy, 1968; Reprinted *Gregorian Semiology*; Abbaye Saint-Pierre de Solesmes: Solesmes, France, 1982.

35. Conklin, D. Discovery of distinctive patterns in music. *Intell. Data Anal.* **2010**, *14*, 547–554. [CrossRef]

36. Moffat, A. Implementing the PPM data compression scheme. *IEEE Trans. Commun.* **1990**, *38*, 1917–1921. [CrossRef]