

# Rhythmic, Melodic and Vertical N-Gram Features as a Means of Studying Symbolic Music Computationally

**McKay, Cory**

cory.mckay@mail.mcgill.ca  
Marianopolis College, Canada

**Cumming, Julie**

julie.cumming@mcgill.ca  
McGill University, Canada

**Fujinaga, Ichiro**

ichiro.fujinaga@mcgill.ca  
McGill University, Canada

Past research has demonstrated that statistical analysis and machine learning can derive meaningful musical insights from features automatically extracted from digitally encoded symbolic music (i.e., notated music, as opposed to audio) in a variety of areas, including composer attribution, regional style and genre (Cuenca / McKay 2021; Cumming / McKay 2021; McKay et al. 2017; Rodríguez-García / McKay 2021). All the aforementioned work was carried out using our open-source jSymbolic2 software (McKay et al. 2018), which extracts 246 features (1,497 dimensions) from symbolic music files. As part of our ongoing work on the next iteration of jSymbolic (jSymbolic3), we have implemented new features based on n-grams. This paper introduces these features and discusses results from repeating previous experiments with the n-gram features added. These experiments involve Renaissance music, but the features themselves are designed to be applicable to any kind of music that can be encoded as a MIDI file.

N-grams can represent diverse types of sequences, and have been used in research ranging from natural language processing to studying DNA sequences. They are well-suited to music, as they can characterize many kinds of musical sequences and transitions. For example, the melodic 3-gram [4,3,5] indicates semitone counts outlining a 4-note ascending arpeggiated major chord in a transposition-invariant way. Researchers including Schubert (2007), Ogiwara & Li (2008), Arthur (2018) and Sears & Forrest (2021) have successfully used n-grams (and related skip-grams and flex-grams) to study music, manually and with automated techniques. We complement earlier approaches by using n-grams in a novel way: studying features extracted from sets of n-grams aggregated from a piece or section, rather than analyzing n-grams directly. These features include statistical measures of which types of n-grams are most common, how dominant the most common ones are and how much diversity there is in the total set of n-grams present. Some n-gram features capture the music as a whole, while others consider only the lowest or highest melodic line.

More broadly, both manual musicological methods and alternative computational tools for studying music, such as Humdrum

(Huron 2002), music21 (Cuthbert et al. 2011), MIDI Toolbox (Eerola / Toivainen 2004) or VIS (Antila / Cumming 2014), tend to focus on local elements, such as melodic themes, chord progressions or contrapuntal modules, and how they recur and vary. jSymbolic's features, in contrast, statistically characterize pieces or sections in their entirety, which allows music to be compared, analyzed or classified in novel ways. The new jSymbolic features extracted from aggregated n-grams permit jSymbolic to maintain this expanded perspective, while at the same time gleaning new kinds of information about local transitions that may have been previously hidden.

The new development version of jSymbolic3 extracts 13, 39 and 24 new features from rhythmic, melodic and vertical interval n-grams, respectively. The experimental results presented in this paper show how combining the new n-gram features with the existing jSymbolic2 features can improve models. To briefly summarize, adding n-gram features to the existing non-n-gram features led to statistically significant increases in classification accuracy in two of three experiments (genre and regional style), and to a small but not statistically significant increase in the other (composer attribution).

A discussion of how individual n-gram features can be identified as characteristic of particular styles is also presented, using analysis based on information gain metrics. This highlights a particular advantage of jSymbolic's bespoke features: they allow users to not only distinguish between musical categories, but also explore which specific properties differentiate the categories statistically, based on musically salient quantities encapsulated by the features. So, for example, features can be used first to train classifiers to propose attributions for anonymous pieces. Then, features extracted from a composer's ground truth exemplars can be analysed to reveal potentially unexpected insights into the composer's style, which may in turn help explain suggested attributions.

This paper will conclude with an acknowledgement of how our research with musical features over the past two decades has benefited greatly from collaborations between musicologists, theorists, computer scientists, music librarians, cognitive scientists and music information retrieval researchers. We hope that our work with n-grams will open up new multidisciplinary collaboration with non-music researchers who have valuable experience utilizing n-grams in their own fields, and that it will encourage more music researchers to incorporate n-grams in their own work. jSymbolic is based on an open-source modular feature plug-in architecture, so in addition to being a research tool, it can also serve as a platform for developing, sharing and applying feature implementations collaboratively.

## Bibliography

**Antila, Christopher / Cumming, Julie** (2014): "The VIS framework: Analyzing counterpoint in large datasets", in: *Proceedings of the 15th International Society for Music Information Retrieval Conference*, Taipei, Taiwan, October 2014: 71–76.

**Arthur, Claire** (2018): "Mass analysis: Renaissance theory versus practice in Palestrina's sacred works", in: *Program and Abstracts of the New England Conference of Music Theorists*, Waltham, USA, April 2018: 6.

**Cuenca, María Elena / McKay, Cory** (2021): "Exploring musical style in the anonymous and doubtfully attributed mass movements of the Coimbra manuscripts: A statistical and machine learning approach", in: *Journal of New Music Research* 50, 3: 199–219.

**Cumming, Julie / McKay, Cory** (2021): “Using corpus studies to find the origins of the madrigal”, in: *Proceedings of the Future Directions of Music Cognition International Conference*, virtual, March 2021: 38–42.

**Cuthbert, Michael Scott / Ariza, Christopher / Friedland, Lisa** (2011): “Feature extraction and machine learning on symbolic music using the music21 toolkit”, in: *Proceedings of the International Society for Music Information Retrieval Conference*, Miami, USA, October 2011: 387–392.

**Eerola, Tuomas / Toivainen, Petri** (2004): “MIR in Matlab: The MIDI Toolbox”, in: *Proceedings of the International Conference on Music Information Retrieval*, Barcelona, Spain, October 2004: 22–27.

**Huron, David** (2002): “Music information processing using the Humdrum toolkit: Concepts, examples, and lessons”, in: *Computer Music Journal* 26, 2: 11–26.

**McKay, Cory / Cumming, Julie / Fujinaga, Ichiro** (2017): “Characterizing composers using jSymbolic2 features”, in: *Extended Abstracts for the Late-Breaking Demo Session of the 18th International Society for Music Information Retrieval Conference*, Suzhou, China, October 2017.

**McKay, Cory / Cumming, Julie / Fujinaga, Ichiro** (2018): “jSymbolic 2.2: Extracting features from symbolic music for use in musicological and MIR research”, in: *Proceedings of the International Society for Music Information Retrieval Conference*, Paris, France, September 2018: 348–354.

**Ogihara, Mitsunori / Li, Tao** (2008): “N-gram chord profiles for composer style representation,” in: *Proceedings of the International Society for Music Information Retrieval Conference*. Philadelphia, USA, September 2008: 671–676.

**Rodríguez-García, Esperanza / McKay, Cory** (2021): “Composer attribution of Renaissance motets: A case study using statistical features and machine learning”, in: Rodríguez-García, Esperanza / d’Alvarenga, João Pedro (eds): *The Anatomy of Iberian Polyphony around 1500*. Kassel, Germany: Edition Reichenberger 401–438.

**Schubert, Peter** (2007): “Hidden forms in Palestrina’s First Book of Four-Voice Motets”, in: *Journal of the American Musicological Society* 60, 3: 483–556.

**Sears, David. R. W. / Forrest, David** (2021): “Triadic patterns across classical and popular music corpora: Stylistic conventions, or characteristic idioms?”, in: *Journal of Mathematics and Music* 15, 2: 140–153.