Julie E. Cumming

Why Should Musicologists Do Digital Humanities?

What is digital humanities in music? And what does it mean to »do« digital humanities?[1] Digital humanities can mean access to materials on line: images of musical sources, scores, inventories, bibliographic data. It can mean constructing a database to organize information about music. It includes projects that make use of music in searchable symbolic notation. More broadly, it refers to new ways of presenting, describing, analyzing, visualizing, and explaining music, using computers. There are at least three different reasons to do digital humanities: to do a better and faster version of things we already do; to do new kinds of research using lots of data; and to learn to think in new ways. To demonstrate what digital humanities can bring to musicologists I will describe some of the ways I have used digital humanities in my own research.

To do a better and faster version of things we already do. Musicologists have always looked at manuscripts, compiled inventories, and made editions of music. Digital humanities makes it easier to do these things, and sometimes to do a better job of it. The wonderful online repositories of digital images (such as Gallica and the Bayerische Staatsbibliothek[2]) make it easy to look at manuscripts, to zoom in to look at details, and to compare different sources for the same work without leaving your desk. RISM online makes it possible to search hundreds of inventories, sometimes with links to digital images.[3] Online critical editions have a huge potential for scholars and performers: links from the modern score back to original sources, the opportunity to choose the variant readings you prefer, and to format the edition according to your needs.[4] Wonderful as these new resources are, they do not fundamentally change what we do.

1 For a useful survey of how musicologist understand digital humanities, see Charles Inskip and Frans Wiering, »In Their Own Words: Using Text Analysis to Identify Musicologists' Attitudes towards Technology«, *Proceedings of the 16th International Society for Music Information Retrieval Conference* (2015), pp. 455–61. http://ismir2015.uma.es/articles/171_Paper.pdf. (Accessed January 5, 2019.)

2 For a useful discussion of two major digital repositories of music manuscripts, see Sarah Ann Long, »Review: International Image Interoperability Framework (IIIF); Gallica; e-Codices: Virtual Manuscript Library of Switzerland«, *Journal of the American Musicological Society* 71, no. 2 (2018), pp. 561–72.

3 See www.rism.info/publications.html#c36. Series A is completely online: see https://opac.rism.info/metaopac/start.do?View=rism. (Accessed January 5, 2019.)

4 For Gesualdo Online, a collaborative project directed by Philippe Vendrix, see https://ricercar.gesualdo-online.cesr.univ-tours.fr/; for the Du Chemin Chansonniers, directed by Richard Freedman, see http://ricercar.cesr.univ-tours.fr/3-programmes/EMN/duchemin/; for the soon-to-be released Marenzio edition, see the article by Laurent Pugin in thi s issue of *Troja*. (Accessed January 5, 2019.)

To do new kinds of research using lots of data. Sorting and organizing information, creating tables, and counting things are traditional scholarly activities. But once there are more than a critical number of sources, variants, or occurrences, it is impossible for a single scholar to control data using traditional methods. Relational databases that include metadata about musical works and sources make it possible to collect multiple different kinds of metadata about pieces and sources and discover new things with queries that explore relationships among different kinds of data.

Searchable symbolic music notation formats (such as MIDI, musicXML, and *kern) make it possible to find and count all kinds of musical events in ways that have never before been possible. For an individual to count the vertical thirds in a single piece of music would take hours (and it is almost impossible to avoid mistakes); music analysis software can search for that same piece of musical data over hundreds or thousands of pieces in minutes, and record and visualize the information.

To learn to think in new ways. When working with computers scholars must know exactly what they want (or mean). For example: when searching for vertical thirds between voices, do we want them between all pairs of voices? One particular pair of voices? Do we want only simple thirds, or do we include compound thirds (10ths, 17ths)? Do we want only thirds that come on strong beats, or that last longer than a particular note duration? Having to make these decisions forces scholars to be extremely precise; this precision can lead to new insights.

To illustrate these points I will share some of my own experiences with digital humanities.

I. Relational databases

My first digital humanities work involved organizing and sorting information using relational databases. In my book, *The Motet in the Age of Du Fay*,[5] I set out to look at how the motet changed between c. 1400 and 1474. In order to do that I had to decide which pieces qualified as motets. I took as my model the work of Alastair Fowler, a specialist in English literature. Fowler describes the novel as follows:

> Turning to prose, we find the status of subgenres … enhanced …. »The novel« has assimilated other kinds of prose fiction. A genre so comprehensive can have but a weak unitary force. Indeed the novel has largely ceased to

5 Julie E. Cumming, *The Motet in the Age of Du Fay* (Cambridge, 1999).

> function as a kind [genre] in the ordinary way.[6] … [But] the novel is still a kind, even if one badly in need of subdivision.[7]

I adapted Fowler's approach to the fifteenth-century motet.[8] I located most of the motets found in manuscripts copied between 1400 and 1475, and assigned each to a subgenre (sometimes to more than one). I used a relational database (Paradox) with the following linked tables: pieces, composers (with dates), manuscripts (with *sigla* and dates), modern editions, and subgenres.

This allowed me to use queries to generate lists and tables, a prominent feature of the book. Table 7.1 (p. 149), for example, lists the subgenres found in Bologna Q15 and shows how many examples of each of those subgenres were represented in other sources of the period. Table 9.1 (pp. 187–9) lists all the English cantilenas found in the Trent Codices and Modena X.1.11, including composers and all the concordant sources for each motet. The appendix of »Widely disseminated motets« (pp. 304–305) lists all the motets from the period with four or more sources and provides the number of sources for each one. A surprising finding was that 17 of the 27 motets with four or more sources (63%) are English. My »Index of works« at the end of the book (pp. 384–99) includes the title, composer, subgenre(s), concordant sources, and modern editions of all the works discussed in the book. All these tables were generated using the database software; after assembling the information in the database, I never had to retype a list of pieces or manuscript sigla – I just generated a new table.

My next project focused on the first five printed books of motets: Petrucci's *Motetti A, B, C, libro quarto,* and *a cinque*), published 1502–1505 and 1508.[9] These motet collections, while published in the early sixteenth century, provide a useful sample of the many different kinds of motets composed at the end of the fifteenth century. Using Microsoft Access database software, I collected detailed information on concordant sources and on imitative texture, resulting in articles about the development of imitative textures and the cultural impact of the prints.

My colleague Peter Schubert's ground-breaking article, »Hidden Forms in Palestrina's First Book of Four-Voice Motets«,[10] provided a way to talk about imitative textures with a great deal of precision, by categorizing different »presentation types« of repeated contrapuntal combinations, or modules. I decided to

6 Alastair Fowler, *Kinds of Literature: An Introduction to the Theory of Genres and Modes* (Cambridge, Mass., 1982), p. 118.
7 Ibid., p. 120.
8 Cumming, Motet (cf. fn. 5), pp. 7–9.
9 RISM 1502^{1} (and 2^{nd} ed., $[1505^{7}]$), 1503^{1}, 1504^{1}, 1505^{2}, and 1508^{1}.
10 Peter N. Schubert, »Hidden Forms in Palestrina's First Book of Four-Voice Motets,« *Journal of the American Musicological Society* 60, no. 3 (2007), pp. 483–556.

use his methodology to understand the kinds of imitative textures in the Petrucci motets. I worked with a team of McGill students (Alexis Luko, Catherine Motuz, Alison Kranias, Adalyat Issiyeva, and Michel Vallières) to collect information about presentation types and time and pitch intervals of imitation at the beginning of each *pars* of each of the motets in the five books (174 pieces, with 355 partes). I later worked with another team of students (Remi Chiu, Jane Hatter, Daniel Donnelly, and Edward Melson) to collect information on text-setting for the same passages.[11] This resulted in my paper from 2012, »Text Setting and Imitative Technique in Petrucci's First Five Motet Prints.«[12] My database allowed me to look at the extent to which imitative presentation types corresponded to the type of text setting. I was able to conclude that there was a strong correlation, as I explained in the conclusion of the article:

> As we have seen, over the last quarter of the fifteenth century nonimitative texture and free imitation (combined with melismatic text setting) gradually gave way to a new style resulting from the influence of the chanson, the introduction of syllabic homorhythm, the use of repeated notes, and the syllabic *soggetto*. Here in the Petrucci motet prints we already find the basic elements of sixteenth-century polyphony, in which the repetition of contrapuntal modules is co-ordinated with syllabic text setting to provide clear and memorable *soggetti*.[13]

Peter Schubert and I were doing research into improvised vocal counterpoint in the Renaissance, with a focus on improvisable canon after one time unit, or »stretto *fuga*,« as John Milsom calls it.[14] We found that these improvisable patterns turned up over and over in composed Renaissance music.[15] We were then

11 This research was funded by the Social Sciences and Humanities Research Council of Canada.

12 Julie E. Cumming, »Text Setting and Imitative Technique in Petrucci's First Five Motet Prints,« *The Motet around 1500: On the Relationship of Imitation and Text Treatment?*. Collection »Epitome Musical,« Centre d'études Supérieures de La Renaissance, ed. Thomas Schmidt-Beste (Turnhout, 2012), pp. 63–90.

13 Ibid., p. 90.

14 John Milsom first used the term stretto *fuga* in his article »›Imitatio,‹ ›Intertextuality,‹ and Early Music,« *Citation and Authority in Medieval and Renaissance Musical Culture: Learning from the Learned*, ed. Suzannah Clark and Elizabeth Eva Leach (Woodbridge, 2005), pp. 141–51. Peter Schubert has produced Youtube videos on improvising stretto fuga: Improvising a canon #1: at the 5th above.mp4, https://www.youtube.com/watch?v=n01J393WpKk. (Accessed January 5, 2019.)

15 See Peter Schubert, »From Improvisation to Composition: Three Sixteenth-Century Case Studies,« *Improvising Early Music: The History of Musical Improvisation from the Late Middle Ages to the Early Baroque*. Collected Writings of The Orpheus Institute 11 (Leuven, 2014), pp. 93–130; and Julie E. Cumming, »From Two-Part Framework to Movable Module,« *Medieval Music in Practice: Studies in Honor of Richard Crocker*. Miscellanea 8, ed. Judith Ann Peraino (Middleton, WI., 2013), pp. 175–214.

asked to write an article on imitation for the *Cambridge History of Fifteenth-Century Music*.[16]

We looked at imitative technique across the fifteenth century, finding the earliest examples of stretto *fuga* in the works of Du Fay in the 1420s. We realized that the information collected in the Petrucci database would allow us to search for evidence of improvisable patterns. Stretto *fuga* for more than two voices by definition has the same time unit between entries; this is the presentation type Peter Schubert calls »Periodic entries.«[17] We therefore searched the Petrucci database for examples of Periodic entries at fairly short time intervals, and sorted them by pitch interval of imitation (see Table 1, generated from the Petrucci database).

Table 12.3 *Four-voice stretto* fuga *in points of imitation at the beginnings of* partes *in the first five Petrucci motet prints (15 examples)*
Underlined scale-degree and pitch-interval patterns occur more than once; asterisk (*) indicates that the soggetto is slightly varied

Scale degrees	Pitch interval of imitation	Voices in order of entry	Time interval in semibreves	Petrucci no.	Composer	Title
Conventional stretto *fuga* using invertible counterpoint at the 12th (13)						
5511	+8; −12; +8	T; S; B; A*	2; 2; 2	104	Pinarol	Surge propera amica mea
5511	+8; −12; +8	Ctb; S; B; Cta	1; 1; 1	509	Obrecht	Laudemus nunc
5511	+8; −12; +8	T; S*; T; B	2; 2; 2	304.3	Josquin	Factum est autem cum
5511	+8; −12; +8	T; S; B; A	2; 2; 2	304	Josquin	Factum est autem cum
5511	+8; -12; +8	T; S; B; A	1; 1; 1	325	Anon.	Confitemini domino
5511	+8; −12; +8	T; S; B; A	2; 2; 2	229	Anon.	Hec est illa dulcis rosa
5511	+8; −12; +8	T; S; B; A	2; 2; 2	431	Anon.	O claviger regni celorum (Ex. 12.15)
1155	−8; +12; −8	A; B; S*; T*	1; 1; 1	227	Anon.	Sancta Maria (Ex. 12.16)
1155	−8; +12; −8	A; B; S; T	2; 2; 2	303.3	Josquin	Liber generationis
1155	−8; +12; −8	A; B; S; T	4; 4; 4	439.5	Josquin	Vultum tuum
1155	−8; +5; −8	S; T; A; B	3; 3; 3	436	Brumel	Conceptus hodiernus
1115	+8; −8; +5	B; S; T; A	1; 1; 1	133.2	Ghiselin	Anima mea liquefacta est
1552	+12; −8; +5	B; S; T; A	1; 1; 1	404.3	La Rue	Salve regina (Ex. 12.19)

Table 1. Cumming and Schubert, »The Origins of Pervasive Imitation« (cf. fn. 16), p. 16.

We were able to show that the most common recurring imitative patterns for the Petrucci motets used four-voice improvisable stretto *fuga* (with invertible counterpoint at the twelfth, resulting from the alternation of octaves and twelfths in the pitch intervals of imitation). This allowed us to demonstrate that imitative techniques came out of an improvisatory practice.

I also worked with Paul Yachnin (McGill, English) on a large team grant entitled »Making Publics, 1500–1700: Media, Markets, and Association in Early Modern Europe.«[18] This project focused on the development of new forms of

16 Julie E. Cumming and Peter Schubert, »The Origins of Pervasive Imitation,« in *The Cambridge History of Fifteenth-Century Music*, ed. Anna Maria Busse Berger and Jesse Rodin (Cambridge, 2015), pp. 200–228.

17 Schubert, »Hidden Forms« (cf. fn. 10), pp. 488–9, pp. 498–504.

18 »Making Publics: Media, Markets and Association in Early Modern Europe, 1500–1700« was supported by a Major Collaborative Research Initiative of the Social Sciences and Humanities Research Council of Canada; see www.makingpublics.mcgill.ca. (Accessed January 5, 2019.)

association around cultural projects and science in the early modern period. I decided to look at the impact of the Petrucci motet prints: what was the »public« for the music in the prints?[19] Once again I was able to use my database, this time to understand a cultural issue.

In addition to the table listing all the pieces in the motet prints, I also had a table in the database of all the concordant sources for all the motets, with information on provenance, format (size and layout, print or manuscript, and partbooks or choirbook), contents (what other kinds of pieces were found in the sources), and dates. Information of this kind can tell us a great deal about the creators and users of the sources.[20] I chose to focus on the sources after Petrucci containing two or more of the Petrucci motets, since this suggests that the compilers of these later sources had more than a passing interest in the kinds of pieces found in Petrucci. I then sorted the sources by provenance and format, as shown in Table 2.

Table 6.4 Later Manuscripts and Prints Including More Than One Piece Found in the First Five Petrucci Motet Prints (for Details, See Appendix)

(a) Italian sacred choirbooks (8; 3 include copies from Petrucci)

(b) German mixed sacred choirbooks (4)

(c) Other choirbooks (7, from the Netherlands, Spain, Eastern Europe, and France; 1 Spanish copy of Petrucci)

(d) French *chansons* (6)

(e) Italian *frottole* and *laude* (4; 1 includes copies from Petrucci)

(f) Sacred partbooks, manuscript and print (11; 2 include copies from Petrucci))

(g) Tablature (3; 1 includes copies from Petrucci)

(h) Music theory treatises (5; 4 include copies from Petrucci)

(i) MSS associated with Glareanus (2; both include copies from Petrucci)

Table 2. Cumming, »Petrucci's Publics« (cf. fn. 19), p. 107.

19 Julie E. Cumming, »Petrucci's Publics for the First Motet Prints,« *Making Publics in Early Modern Europe: People, Things, Forms of Knowledge*. Routledge Studies in Renaissance Literature and Culture 13, ed. Paul Edward Yachnin and Bronwen Wilson (New York, 2010), pp. 96–122.

20 See Julie E. Cumming, »Sources and Identity: Composers and Singers in Darnton's Communications Circuit,« *Sources of Identity: Makers, Owners and Users of Music Sources Before 1600*, ed. Tim Shepherd and Lisa Colton (Turnhout, 2017), pp. 25–38. For another study on the origins of the repertoire in the Petrucci motet prints, see Julie E. Cumming, »From Chapel Choirbook to Print Partbook and Back Again,« *Cappelle musicali fra corte, stato e chiesa nell'Italia del rinascimento. Atti del Convegno internazionale Camaiore, 21–23 ottobre 2005*. Historiae musicae cultores 108, ed. Franco Piperno, Gabriella Biagi Ravenni, and Andrea Chegai (Florence, 2007), pp. 373–403.

What Table 2 tells us is that the Petrucci motets were found in sources all over Europe. Categories (a)–(c), large choirbooks, indicate that the motets were performed in ecclesiastical institutions. Categories (d)–(f), containing secular music and partbooks, suggest that the motets were sung in domestic situations and confraternities. Category (g) indicates that the motets were arranged for instruments; and (h)–(i) demonstrate how important the motets were for music theorists.[21] My database provided a way to show the range and diversity of these »publics« for the Petrucci motets.

II. Learning to think in new ways with searchable symbolic notation, using more precise definitions of terms and concepts

The work on imitation I did with the Petrucci database required people to make observations about the music and record them in a database. This turned out to be very effective, but it is also subject to errors, since it is difficult to be completely consistent when looking at music. My McGill colleague Ichiro Fujinaga, a professor in the Music Technology Area of the Schulich School of Music, inspired me to start thinking about using searchable symbolic notation and computer analysis tools. Peter Schubert and I wanted to be able to search for repeated contrapuntal combinations (modules) in Renaissance music, so I applied for and received a grant to fund the development of this kind of tool: »ELVIS: Electronic Locator of Vertical Interval Successions. The first large data-driven research project on musical style.«[22]

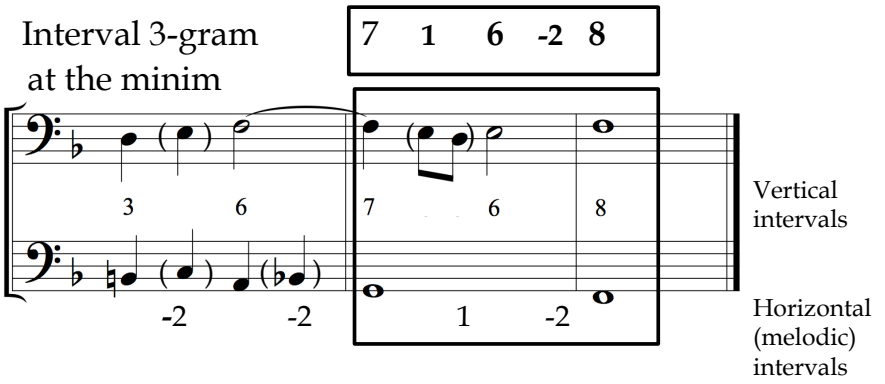
We needed to compile a substantial corpus of music in searchable symbolic notation, so we created the ELVIS database: an online database of searchable scores of polyphonic music 1300–1900, taken from a wide variety of sources, including online repositories and donations of transcriptions from individual scholars. We created a software tool (VIS, for »vertical interval successions«) that can search for repeated contrapuntal patterns; and we did research on musical style using the ELVIS data and tools. Our work took different forms – from a large-scale study

21 See Cristle Collins Judd, *Reading Renaissance Music Theory: Hearing with the Eyes*. Cambridge Studies in Music Theory and Analysis 14 (Cambridge, 2000) for a full discussion of this issue.

22 I was the principal investigator of this Digging into Data Challenge grant, which had an international team. The Canadian co-investigators were Peter Schubert, Ichiro Fujinaga, Jonathan Wild, René Rusch, and Cynthia Leive; we were funded by the Social Sciences and Humanities Council of Canada. Also on the grant, but working on different projects, were Michael Scott Cuthbert and Ian Quinn (USA), Frauke Jürgensen and George Coghill (Scotland). The VIS analysis framework uses Cuthbert and Ariza's music21 (http://web.mit.edu/music21/) as a back end. There is a web application that is not functional right now, and a VIS API. The database and the software can be found at http://elvisproject.ca/. (Accessed January 5, 2019.)

of style change over time, to a focused examination of a small corpus of duos.

VIS was built to focus on interval successions between pairs of voices in polyphonic music. We wanted to find all the occurrences of chunks of counterpoint, defined as successions of vertical intervals linked by horizontal (or melodic) intervals. We looked at interval n-grams (or contrapuntal modules) of different lengths, where n = the number of vertical intervals (see Example 1). An n-gram that occurs more than once in a composition is a »module«: a chunk of counterpoint that is repeated. Our notation for n-grams alternates numbers representing diatonic vertical intervals (7, 6, 8 in the boxed 3-gram in Ex. 1) with numbers representing the horizontal (or melodic) intervals of the lower voice: 1 (unison), -2 (descending second) in the boxed 3-gram. The intervals of the upper voice (which would be -2, +2 for notes at the beginning of the minim in the boxed 3-gram) are not necessary, since they result from the other two sets of numbers. This provides a convenient shorthand notation for a segment of two-voice counterpoint.



Example 1. Interval 3-gram, showing cadential suspension.

When you work with a computer you have to make numerous decisions about exactly what you are looking for. Our decisions affect the results in significant ways.[23]

• What is the sampling rate? Do we look at every attack, every semiminim, minim, semibreve, etc.?

23 We experienced this problem first hand, when we asked two different people to find 3-grams in duos by Josquin and La Rue. They kept getting different values for the occurrences of each 3-gram, and for the total number of 3-grams, although they were using the same corpus. We discovered that they were defining the 3-grams differently; in particular, sustained notes and rests were not defined the same way.

We chose to sample at the minim (i.e. we sampled only vertical intervals at the beginning of each minim). This value corresponds to the basic level of counterpoint in most Renaissance music, and captures cadential suspensions.[24] This meant that we ignored notes (such as passing semiminims or fusae) between each minim (see notes shown in parentheses in Example 1).

• How do you handle sustained notes or repeated notes in both voices?

If you sample at every minim, whenever you have a sustained note in both voices longer than a minim, you will get a 2-gram with no contrapuntal motion at all. We therefore chose to eliminate repeated or sustained notes in both voices after the first minim; the next vertical interval was sampled when one voice changed pitch.

• How do you deal with rests? Can an n-gram extend across a rest?

Because the vertical interval is central to our concept of n-gram, we said that there had to be two voices sounding for each interval in our n-gram. We therefore ended our n-grams on the last sounding minim before a rest in one or both voices; we then started a new string of n-gram after the rest.

• How long should »n« be?

We looked at various values for n, from 2 to 10, but ended up focusing on 3-grams in most of our work. 3-grams are long enough to capture cadential suspensions, but short enough so that there is lots of data (many recurrences of most 3-grams). The longer the value of n, the fewer examples there will be of the n-gram.

One of the major motivations for the ELVIS project was to describe style change using specific data, not just subjective impressions. We therefore decided to look at three-gram distribution over the course of the Renaissance. We assembled test sets for three style periods, named after representative composers for each period: Ockeghem (1440–85), Josquin (1485–1521), and Palestrina (1540–85).[25] We chose to visualize our findings with a figure that combines a Venn diagram, a »3-gram cloud« in which the size of the 3-gram indicates its relative frequency, and a timeline that moves around the diagram from left to right (Figure 1). The diagram includes only repeated interval 3-grams (modules) that constitute greater than 0.2% of the 3-grams in at least one of the test sets.

24 It is also possible to choose different values, depending on the musical context or style. Alexander Morgan, a McGill graduate student who worked with Peter Schubert and me, developed a tool that dynamically finds the appropriate note value to sample for contrapuntal analysis. See his dissertation, »Renaissance Interval-Succession Theory: Treatises and Analysis« (Ph.D dissertation, McGill University Libraries, 2017), http://digitool.Library.McGill.CA/R/?func=dbin-jump-full&object_id=145547. (Accessed January 5, 2019.)

25 Christopher Antila and Julie Cumming, »The VIS Framework: Analyzing Counterpoint in Large Datasets,« *Proceedings of the 15th International Society for Music Information Retrieval Conference* (2014), pp. 71–76: p. 73. http://www.terasoft.com.tw/conf/ismir2014/proceedings/T014_162_Paper.pdf. (Accessed January 5, 2019.)

Most visually striking is the prominence of the cadential suspension, »7 1 6 -2 8.« It is the most common 3-gram in the dataset, and it is shared by all three style periods (as are two other 3-grams that end with the »7 1 6« suspension). The diagram also shows evidence of stylistic change. Most notably, the Josquin and Palestrina test sets show a higher level of repetition than the Ockeghem set. The number of repeated 3-grams is higher in the Josquin test set (with seventeen 3-grams) than either the Ockeghem or Palestrina sets (both with eleven 3-grams). These data suggest an increase in repetition of contrapuntal modules from the Ockeghem to the Josquin generations; most of the repeated 3-grams (modules) in the Josquin generation were retained in the Palestrina generation. Descending parallel 10ths (10 -2 10 -2 10) are especially prominent in the Josquin test set.[26] This approach to describing style change has the potential to work for almost any repertoire.

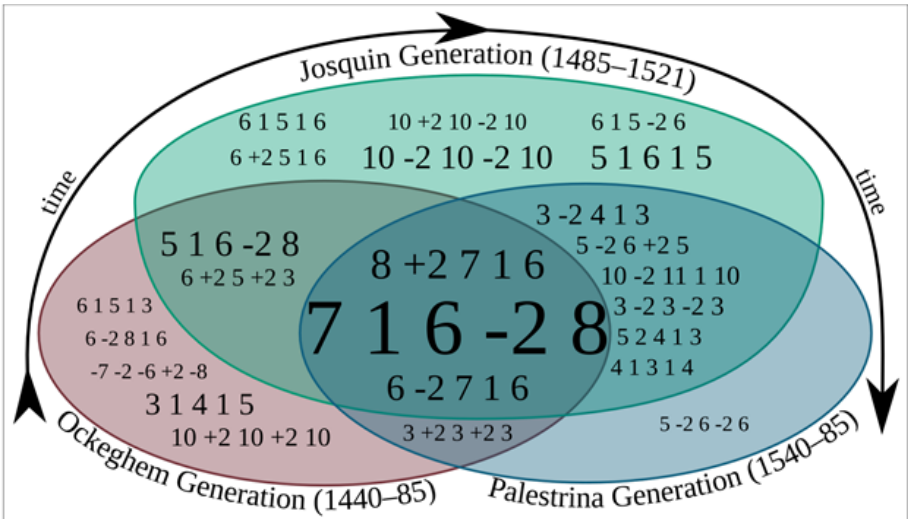


Figure 1. Hybrid Venn diagram, 3-gram cloud, and timeline, from Antila and Cumming, »The VIS Framework« (fn. 26), p. 74.

In another study, Peter Schubert and I looked at contrapuntal repetition of 3-grams in the Lassus duos of 1577.[27]

26 These findings correspond to my claims in Julie E. Cumming, »From Variety to Repetition: The Birth of Imitative Polyphony,« *Yearbook of the Alamire Foundation* 6, ed. Bruno Bouckaert, Eugeen Schreurs, and Ivan Asselman (Peer, Belgium, 2008), pp. 21–44.

27 Orlando di Lasso, *Novae Aliquot … ad duas voces cantiones suavissimae* (Munich, 1577). RISM 1577c (B/I) = L 902 (A/I). For a modern edition see Orlando di Lasso, *The Complete Motets*, vol. 11. Recent Researches in the Music of the Renaissance 103, ed. Peter Bergquist (Madison, WI,

The amount of musical repetition is a style feature that can be used to distinguish pieces, composers, style periods, and genres. In order to determine whether one composer or piece uses more repetition than another, we have to quantify the repetition. The best way to do this is with a computer, since it is difficult for a human to count every repeated musical item accurately.

So what do we mean by repetition? And what exactly do we count? As I sat down to figure this out, I realized that there are at least three different ways to quantify repetition: length of repeated things; number of repetitions of each individual item; number of different things that are repeated. As a test case, we decided to look at all three kinds of repetition in the Lassus duos. We posed the questions as follows.

1. How long are the repeated n-grams? Or: What is the longest n-gram that repeats?
2. How many times do n-grams repeat? Or: What is the largest number of repetitions of any n-gram?
3. How many different repeated n-grams (modules) are there?

We then added the scores for each type of repetition for each duo, to get a total score.

		1. How long?	2. How many times?	3. How many different modules?	Sum of 1, 2, & 3
1	Beatus vir	6	3	6	15
2	**Beatus homo**	3	3	3	9
3	Oculus non vidit	8	2	9	19
4	Justus cor suum	10	2	9	21
5	Expectatio justorum	10	3	4	17
6	Qui sequitur me	4	2	8	14
7	Justi tulerunt	4	2	5	11
8	**Sancti mei**	7	5	10	22
9	Qui vult venire	4	3	9	16
10	Serve bone	4	2	6	12
11	Fulgebunt justi	5	2	8	15
12	Sicut rosa	5	2	8	15

Table 3. Quantifying three different types of repetition in the Lassus vocal duos of 1577. Cells with thick borders contain the highest values in each column; grey cells contain the lowest.

1995). We studied the twelve vocal duos; we did not look at the twelve untexted duos in the book. The initial version of this project, which I discuss here, focused on quantifying contrapuntal repetition in the Lassus duos. It was presented at Med-Ren Certaldo, July 2013: Julie Cumming and Peter Schubert, »Another Lesson from Lassus: Quantifying Contrapuntal Repetition in the duos of 2017,« presented at Med-Ren Certaldo, July 2013. The published version of this paper went in a different direction: see Peter Schubert and Julie Cumming, »Another Lesson from Lassus: Using Computers to Analyse Counterpoint,« *Early Music* 43, no. 4 (2015), pp. 577–86.

Admittedly, the results are not easy to make sense of. We can see that there is a wide variety of values for the different types of repetition across the collection of duos. One duo, no. 8, has a high score of 22 in the final column, while the lowest score in that column, for duo no. 2, is only 9. Even in a very controlled data set like the Lassus duos there can be a wide range of values for different kinds of repetition. But more important than the results for this small set of pieces is the way it caused me to think about repetition. Before I began this project I thought I knew what repetition was; now I see it is a complex problem. Using the computer to study repetition in the Lassus duos forced us to clarify, refine, and expand our definition of repetition. Which type of repetition do we want to quantify? Are different types relevant for different styles of music? Should different types of repetition be weighted for different repertoires?

III. Conclusion

Digital humanities, therefore, has a great deal of potential for musicology and musicologists. In an age where more and more information is available, we need tools that will help us organize, search, compare, and query that information. Now that all scores are created by means of notation editors, we have access to lots of repertoire in searchable symbolic notation. We need to begin to explore its potential, to move beyond general impressions about style and style change, and to use real data to back up our claims.

There are of course significant challenges. How do you find people who can code and read music? How can we retrain in middle age? Digital humanities normally involve working in teams: a musicologist and a programmer, a group of students gathering data or doing transcriptions, a group of colleagues. This is a challenge in a relatively traditional field where the single-author paper or monograph are the principal currency for hiring, tenure, and promotion. There are signs of change, however: granting agencies (based on a science model) are enthusiastic about digital projects and about research teams. Young people are engaged with technology and eager to explore. In a shrinking job market, digital humanities can lead to jobs outside the academy. My experience has proved to me that working with teams of students and colleagues on digital projects makes it possible for all of us to do important research, and teaches us all to think in new ways.