ELVIS

Electronic Locator of Vertical Interval Successions:
The first large data-driven research project on musical style

Canada (McGill): Julie E. Cumming, PI (for Canada, and for the whole project). Ichiro Fujinaga, Cynthia Leive, René Rusch, Peter Schubert, Jonathan Wild.
US: Michael Cuthbert (MIT) PI; Ian Quinn (Yale).
UK (U of Aberdeen): Frauke Jürgensen, PI; George Coghill.

- Overview of goals and objectives

The McGill team has focused on developing a searchable database of music in symbolic notation, as well as search and analysis software freely available online and easily accessible to users with minimal computer skills. The ELVIS database contains more than 6,000 pieces of music in symbolic notation, searchable by date, composer, type of piece, number of voices, instrumentation, etc.; the files can be downloaded and used in the software. Our music files come from collections freely available online and from the generous donation of many other music scholars. Our software, VIS (for Vertical Interval Successions), can search and represent data about the relative frequency of contrapuntal patterns (vertical intervals linked by melodic motions) in one piece or thousands of pieces. VIS is based on `music21` (see the US team report below), but has a graphical user interface that is easy for anyone to use. It will be available as a web-based application (no need for installations on different operating systems), and will also incorporate graphing and statistical capabilities and multiple types of queries. We have also worked on two specific research questions: describing and representing style change in music, and how to quantify contrapuntal repetition in music.

The US team has focused on understanding the harmonic progressions underlying eighteenth- and nineteenth-century music and on developing software for rapid querying and reducing of musical scores in general. In learning more about harmonic progressions, we have taken both a macroscopic and a microscopic approach. On a large scale, we have sought to learn the most common chord progressions for a period directly from thousands of source scores, yielding many millions of notes. Composers use thousands of different combinations of vertical successions in any piece, but the vast majority of these harmonic moments are elaborations or variations on a dozen or fewer basic underlying chords. What these basic chords are, however, has generally been learned from music theorists and do not, as we have found, necessarily coincide with the usage in actual music. Chords that were thought to be fundamental emerge as too rare and easily subsumed into another category; on the other hand, chords that were thought to be variants on other chords in fact turn out to have their own unique behaviors and are so common as to be classified as among the basic chords. On a smaller scale, the arrangement of notes within chords has also been an object of study, which has also revealed important contradictions with the received norms of behavior taught in music theory classes. Both of these studies, along with studies by the other national ELVIS teams and by unaffiliated researchers, have relied on the underlying software library, `music21`, developed by the US PI and his team at MIT. A focus of their work in ELVIS has been to take the software that existed before the DiD project and improve it so that it scales to the realm of big data. Integrating with cluster and cloud computing, removing bottlenecks that only appear when the data size moves into the gigabyte and terabyte range, and solving problems that arise when the data are partially corrupt or inaccurate but too large to fix by hand (such as the output of hundreds of thousands of printed scores run through optical music recognition software) have all been top priorities for `music21`'s development in ELVIS. Additionally, new

libraries for indexing and retrieving musical data and metadata and for performing automatic reductions of complex scores have been developed.

The UK team has pursued several strands of investigation. First is the extension of previous research on finding repeated contrapuntal patterns, to solving questions of inexact repetition, using Palestrina's first book of motets as a test set, from which we branch out into the mass settings, and finally the works of other composers. Second, we have invented a metric for the assessment of possible tuning systems for polyphony based on the prevalence of different chord, and we are now expanding this work into the analysis of highly chromatic progressions in the works of late sixteenth-century music by Gesualdo. In a third investigation, we worked with a psychologist studying memory. We collected reactions of subjects to several works of Monteverdi, when they had been primed with other works of Monteverdi, or with works of Palestrina. This was compared with computationally-obtained statistical assessments of dissonance content in the works of the two composers. This study will be published in the *Journal of Interdisciplinary Music Studies*. Finally, we are looking to combine elements of all three projects to study the problem of missing-part completion, using early seventeenth-century English polyphony as a starting point.

- International collaboration across disciplines and domains

The three national teams worked fairly independently. We maintained contact through the ELVIS website, which included minutes for all the McGill meetings, and through frequent e-mail contact, especially with respect to problems and improvements in `music21`. There were two meetings in Montreal where all the teams were represented, and a meeting at an international conference in New Orleans with North American team members and the advisory board. The focus was different in each team, but all three fit into our original conception. The interactions were mutually beneficial. Collaborations between scientists and humanists have been extremely fruitful for both sides: humanists have learned to ask new questions and clarify their definitions, while scientists have relished the computational problems involved in the work.

- Indicators of success; Measuring impact; Knowledge dissemination

Our work has been presented and discussed at numerous international conferences including a packed panel discussion at the American Musicological Society in Fall 2012. ELVIS has garnered a great deal of interest among musicologists and music theorists who had previously been suspicious of quantitative work. The search and analysis capabilities of ELVIS are a central feature of a SSHRC Partnership Grant application going forward this fall (LOI approved), entitled SIMSSA (Single Interface for Music Score Searching and Analysis). We are leaders in computational musicology and music theory.

- Capacity building and training (students and highly qualified personnel)

Student training has been a central part of the ELVIS project, especially at McGill, but also in the Yale component of the US team. We have had a large team of students (two undergraduates, six masters, and eleven doctoral) working on the project. Some are from the humanistic disciplines of Musicology and Music Theory; others are from Mathematics, Computer Science, and Music Technology; several have skills in both areas. The interactions and teamwork among the students, as well as many hours of labour, testing, and retesting have been essential to the project. The ELVIS experience has provided the students with new skills that will make them more employable and more successful in their own work.