Grant Proposal

The Andrew W. Mellon Foundation
Scholarly Communication Program
Proposal date: July 17, 2001
Funding dates: October 2001 - March 2003

Digital Encoding for Medieval Chant Transcription
Phase One Project
Principal Investigator: Prof. Thomas Forrest Kelly
Chair, Department of Music, Harvard University
e-mail: tkelly@fas.harvard.edu

ABSTRACT

The thousands of medieval manuscripts containing musical notation are among the monuments of Western culture. Their study, over many generations, has allowed scholars to retrace the foundations and the development of Western music and have contributed to many scholarly disciplines. These documents are best studied comparatively, and much work has been done in this area. But comparison is limited by the varieties and styles of musical notation which, although they have basic conceptual elements in common, vary enormously in detail, and as to what information is transmitted.

This project seeks to establish a uniform means of providing comparable transcriptions of the musical content of these documents. Since many of the earliest notations (9th-11th centuries) provide indications of melodic direction but not of the pitches of individual notes, transcription in modern musical notation of early documents is often speculative with respect to pitch; it necessarily omits much information as to nuance, performance, and rhythm which the earlier notations provide.

Our goal is to allow scholars to record, in machine-readable and searchable form, the enormous corpus of medieval music (which includes not only chant but also a wide range of other genres, sacred and secular). By this facility, future transcriptions can be compared and studied with little or no loss of information. The detailed comparative study that such databases will eventually allow will permit highly important insights into the nature and development of medieval music which are not possible at present.

CONTENTS

Funding Request & Personnel ................................................................. 2
Overview ................................................................................................. 2
Explanation of the Challenges ................................................................. 4
Previous Work ......................................................................................... 5
Phase One Goals & Procedures .............................................................. 6
Future Directions ..................................................................................... 7
Management Plan .................................................................................. 8
Budget ..................................................................................................... 8
Appendix A: References Cited ................................................................. 9
Appendix B: Curricula Vitae of Principal Personnel ................................. 10
Appendix C: Samples from Neumed Documents ................................. 12
Appendix D: Glossary of Musicological Terms ................................. 17
FUNDING REQUEST & PERSONNEL
Institution to be funded. This grant request is being made by Harvard University, Department of Music.

Funding requested. Total of $150,000 for 18 months (Oct. 2001-Mar. 2003) for project Phase One.

Principal personnel. (see curricula vitae in Appendix A)
Principal Investigator (P.I.): Prof. Thomas Forrest Kelly; Chair, Department of Music, Harvard University
Software Engineer: Prof. Louis W.G. Barton; University of Oxford (external); contract programmer

Advisory Board.
Prof. Thomas Forrest Kelly (Moderator); Department of Music, Harvard University
Prof. Louis W.G. Barton; Software Engineering Programme, University of Oxford (external)
Prof. John A. Caldwell; Faculty of Music, University of Oxford
Prof. Ugo O. Gagliardi; Division of Engineering and Applied Sciences, Harvard University
Prof. James Grierson; Faculty of Music, University of Western Ontario
Prof. Andreas Haug; Center for Medieval Studies, NTNU, Trondheim (Norway)
Dr. Peter G. Jeavons; Computing Laboratory, University of Oxford
Prof. Bradford Maiani; Department of Music, University of North Carolina at Chapel Hill

International collaborators.
Prof. John A. Caldwell (head of team); Faculty of Music, University of Oxford
Dr. Emma Hornby; Faculty of Music, University of Oxford
Dr. Ruth Ripley; Department of Statistics, University of Oxford

Consultative partners.
Dr. Julia Craig-McFeely; DIAMM (Digital Image Archive of Medieval Music)
Dr. Annalisa Doneda; Scuola di Paleografia e Filologia Musicale, Università di Pavia (Italy)
Dr. Consuelo Dutschke; Digital Scriptorium Project, Rare Book & Manuscript Library, Columbia Univ.
Dr. Charles Faulhaber; Digital Scriptorium Project, The Bancroft Library, Univ. California at Berkeley

OVERVIEW
Mission. Digital Encoding for Medieval Chant Transcription (“the Project”) will produce a viable, long-term solution to an important problem in content access, cultural conservation, and scientific study of medieval chant manuscripts. Specifically, the Project will develop a much-needed software infrastructure that will enable digital transcription of the archaic forms of musical notation that were used in these manuscripts. Transcriptions created with this software will be interoperable across computer platforms, serve a variety of uses, and have long-term usefulness. Anticipated uses include public access to content, programmable analysis and comparison of chants, easier creation of examples in books and articles, interactive training of students, and so on. This interdisciplinary project brings together computer scientists and some of the world’s leading authorities on medieval music notation. An importance of this project is as a test case for applying modern software-engineering techniques to a qualitative field at a high level of sophistication.

Extent of domain. A chant manuscript is typically a codex consisting of many individual chant ‘documents’. A chant is a melody and text for one section in the traditional division of the Western liturgy; a chant melody is an integral piece of music, much like a ‘song’ in the modern sense. Chant melodies always appear above a Latin text that was sung [see Appendix C, Sample #1], but many codices include sections of spoken text without melody. Scholars estimate that millions of chant documents likely have survived from the period of about A.D. 800 (when music writing in the West was likely invented) to about 1550 (when mechanical printing of music became feasible), but no one yet knows how many of these documents exist. In the late medieval period, secular music was also written down; nevertheless, it is liturgical chant that makes up the great bulk of music that has survived from this period.
Access. Today, chant codices are widely dispersed in rare-book libraries in various parts of the world. Due to their fragility and irreplaceability, libraries are increasingly setting strict limitations on access to them even by scholars. Since the late 19th century [34], many efforts have been made to increase access through photographic reproduction [1][8][21][33][35]. More recently, initiatives are producing digital images for access via the Web [9][12][15][17]. The second major thrust in ‘computerization’ of these manuscripts is the building of online catalogs of document descriptions [5][10][14][24]; catalog entries may potentially be hyperlinked to online images.

Importance of content. Despite these steps to improve access and to preserve documents as photographs, a very important dimension of the problem has been largely overlooked. Given access to documents, it can be highly difficult for modern readers to decipher the archaic symbols that scribes used for denoting music [see Appendix C, Samples #1 through #5]. The greatest value of these documents is not visual (though some have beautiful illuminations), but semantic. As such, the notion of ‘access’ may be expanded to include access to content. Transcription to modern musical notation does not provide this access as it typically ‘flattens’ the semantics of neume notation (see § “Explanation of the Challenges”). Similarly, the notion of ‘preservation’ can include cultural preservation. Knowledge of how to interpret these culturally important documents has suffered greater deterioration than have the physical artifacts.

Though expert musicologists can decipher most chant symbols, significant progress in their field will depend on much more scientific means of analysis than are now possible. The key will be to use computers for content search, comparison of perhaps thousands of documents, melodic pattern matching, statistical analysis, and so on. Bringing computer power to bear on this large body of documents will vastly increase the possibilities for future research, including accurate dating of manuscripts, tracing the origins of melodies and their lines of transmission, and reconstructing the genesis of musical script.

Much research on medieval chant depends on comparing many manuscripts and considering transmission by means of repertories and variants. Much has been accomplished in this area using verbal texts only (for example, the CANTUS Database [5], and Hesbert’s study of antiphoners), but comparison of musical texts has been hampered by the incommensurability of early pictographic notations and later pitch-specific notations. This project will provide a software infrastructure for searchable databases that will allow the comparison of melodic similarities and variants across a wide spectrum of sources.

As chant documents are among the few primary sources from the Middle Ages, their systematic study may shed indirect light in other scholarly disciplines. Another residual benefit of transcription is that it will help refine the requirements for digital image photographs (see § “Future Directions”).

Timeliness. Support from the Mellon Foundation for Phase One of this project is crucial. It will capitalize on the momentum of preliminary work already done [2], and take advantage of the willingness by several experts to contribute their time to the Project. There is also urgency due to widespread interest in making digital transcriptions [4][11][19][25][29][30][31][38]. Delay in creating a comprehensive data representation will result in fragmented digital libraries [12]. Data representations of convenience (where the main interest is not in the representation itself) are likely to be mutually incompatible. For instance, transcriptions are already being made using the Access™ database program [4][11], which is unusable on non-Microsoft platforms. For expediency, people tend to simplify data representations by tailoring them to specific purposes, but information left out of one representation might well be needed in another. For example, data representations for making printable documents in neume notation [29][31] leave out much information needed for musicological analysis. If this trend is not averted, there will arise a Babel of incompatible ‘standards’, duplication of effort, and investment in transcriptions of limited value.

Sustainability. A ‘universal’ data representation for chant can avert fragmentation of digital libraries and ensure long-term viability of transcription investments. The data representation will be designed in the Private Use Area of the Unicode™ Standard [37] and XML (Extensible Markup Language) [39]. An XML document type definition (DTD) will enforce proper formatting of the XML; a new feature of XML [3] will also enforce proper sequencing of character data. To encourage widespread adoption of the data representation, the Project will develop platform-neutral Java™ classes for reading and writing data in this data representation. These classes will be ‘callable’ from other programs, thus providing a layer of functional abstraction between the data representation and programs that use it.
Extensibility. “Extensibility” allows addition to the data representation should a need arise in the future. If, however, a standard can be modified, then there is risk that older transcription files will become incompatible. With new standards, there is typically some conflict between the needs for extensibility and stability. The sensible solution used by the Unicode® Consortium is (a) to leave gaps in the character sequence for future additions, (b) to allow time for public deliberation before any change is adopted, and (c) not to invalidate any code previously adopted. This Project will adhere to these three principles.

If current scholarly opinion “reclassifies” a particular written symbol from one abstract neume to another, then some transcriptions might be judged in error, yet their data format would still be compatible. A given document might have several alternate transcriptions, perhaps reflecting different transcribers’ interpretations about which abstract neume a written symbol represents. The DTD will allow an XML file to contain the transcriber’s name, transcription date, editorial comments, and so on. Many institutions can produce transcriptions independently, and no central authority is needed for approval of transcriptions.

Collaboration. The Project will be conducted in collaboration with the Oxford Chant Group [28]. They have received funding in the U.K. from the Arts & Humanities Research Board [4]. Their major focus is on algorithms for analysis of digital transcriptions. They have transcribed a few early sources to an in-house data representation, but they are keenly aware of the need for a universal encoding standard. The two teams will cooperate closely on requirements specification for the character data and the DTD.

EXPLANATION OF THE CHALLENGES

Pioneering work. This project is unique in directly confronting foundational questions about data representation for medieval music notation. No plan has ever been proposed for encoding chant symbols in a manner that captures the semantic subtleties of the symbols and that lends itself to a variety of uses. Substantial difficulties thwart any easy solution; these include the huge number of distinct symbols used, scholarly uncertainty about some symbols, and a need to allow for semantic equivalence between symbols that derive from very different graphical paradigms. Resolution of the difficulties requires innovative software engineering and semiotic analysis at the frontier of modern understanding about this writing system. Though total success is not assured, a practical and usable solution will be found. The Latin texts underlying the music are usually known from control sources, and their transcription is not problematic.

Classification. A neume (Greek-Latin ‘sign’ or ‘gesture’) is a notational sign representing a musical gesture [6]. Such signs may represent a single note or a group of notes. In their early forms they indicate melodic direction without specifying the pitches involved [18]. By the Renaissance, square-neume notation with staff lines was in widespread use [see Appendix C, Samples #6 and #7]. This type was largely normalized by mechanical printing of music; it is recognizably similar to the type of music writing used throughout the world today. From the earliest known neumatic notation (likely 9th century) until the normalization of music writing, there were dozens of distinct regional and historical genres. The distinctions between these genres are more profound than simple differences of handwriting. Considering all these genres, there are thousands of distinct symbols. Scholars theorize that across the Middle Ages there was a common vocabulary of concepts about musical gestures [36]. Thus, the following ten symbols are believed to be a few of the variants a single concept, that of the two-note clivis neume.

Classification of neumes representing more than two notes can become rather complex, particularly because different genres might use a different number of symbols to represent the same neume.

Incompatibility with modern notation. The core reason a new data representation is needed is that neume notation does not correspond at all well with modern musical notation. Incompatibilities include the vast number of neumatic symbols and the concepts scribes used to classify aural phenomena. The neumatic system was more complex than the modern system. Modern notation is ‘quantificational’, in that notes denote specific pitches and durations. Earlier notational genres were concerned foremost with recording subtle nuances of vocalization, for which no equivalent symbols exist in modern notation.
Furthermore, the earlier symbols seem to be pictographs of melodic shape, not individual notes. Force-fitting medieval chant into the modern idiom would impose a *diatonic* and *mensural* interpretation that may be unwarranted and anachronistic, and it loses any trace of the vocal subtleties of neume notation.

**Non-‘alphabetic’**. The Project researchers believe that the basic unit of transcription should be the ‘abstract’ neume, not the written symbol. The notational genre should be identified globally in the XML ‘meta-data’. Transcribing abstract neume characters will allow for efficient comparison of melodies across genres. It is the opinion of these researchers that ‘alphabetic’ transcription of chant (in the sense of a one-to-one correspondence between symbols and data codes) would not satisfy the stated requirements.

**Unheighted neumes**. In early neumed sources, scribes did not write horizontal lines (like modern ‘staff’ lines) as guides for vertical placement of symbols. These so-called ‘unheighted’ neume genres pose added difficulty. A comprehensive data representation must accommodate both heighted and unheighted neumes, especially to allow for melodic search and comparison across notational genres. In anticipation of optical character recognition (see § “Future Directions”), the data representation must allow a metric of vertical placement on the page, so that future scholarship can use this information when present.

**Uncertainty in the data.** Two types of transcription uncertainty should be recorded in the data.

1) Identification: even an expert might be unsure how a symbol should be interpreted. 2) Scribal error: a scribe might have made a mistake, especially with nascent writing styles or when copying another source. For each type, there will likely be four levels of qualitative certainty: “is” (the default value); “probably is”; “might be”; and “is completely uncertain.” Guidelines for transcribers will be drawn up to facilitate consistent usage. Document damage should be recorded in XML markup tags, not in the character data.

### PREVIOUS WORK

#### Existing encoding standards.

There are today well over a dozen standard data representations for music [32], including DARMS (Digital Alternate Representation of Musical Scores) [7], HyTyme (Hypermedia/Time-based Structuring Language) [20], NIFF (Notation Interchange File Format) [27], SMDL (Standard Music Description Language) [26], and so on. Good [14] is developing a representation for music in XML. Among other problems regarding neume notation, all these data representations are strongly structured around modern musical concepts, including time-valued notes, discrete note pitches, and the grouping of notes into regular sequences or ‘measures’. This is quite appropriate for modern pitch music, but ineffective for neumed chant (see § “Explanation of the Challenges”). Roland [30] has proposed an encoding for music symbols in the Unicode™ Standard that would include about a dozen neumatic signs. Regarding neumed transcriptions, his plan is suited for only a subset of square neumes, such as one finds in the modern *Liber Usualis* [23] [see also Appendix C, Samples #6 and 7], and it does not solve the challenges discussed above. McGee and Merkley [25] have worked on computerized pattern recognition for lined, square-neume notation but they are silent on the problem of data representation.

**Coding practice.** The most significant advance in data representation has been the Unicode™ Standard—a result of extraordinary cooperation in the software industry. The Unicode® Consortium has done an exceptionally good job of designing the standard with forethought. This project will adhere as much as possible to the coding guidelines set forth by the consortium in its official publication [37].

**Taxonomy of neumes.** Since the pioneering work of the monks of Solesmes [34], scholars have recognized fundamental similarities in the various Western notational systems. Recent studies, including those of Cardine [6], Stäblein [35], Corbin [8], and Solesmes [33] have done much to systematize and organize current understanding of notational styles, though there remains much to be learned. Certain signs that have performance significance for vocal inflection (like the *oriscus*, the *quilisma*, the *pressus*) are still poorly understood. Current research by Wulf Arlt [1], Charles Atkinson, Kenneth Levy [22], Andreas Haug, and others, seeks to understand the conceptual basis of neumatic notation and the nature of its origins. The neume taxonomy developed by this project will build upon the work of these scholars.

**Feasibility.** During the past several years, Louis Barton [2] has been studying the feasibility of a comprehensive data representation for encoding neumed manuscripts, including partial solution to several problems discussed above in § “Explanation of the Challenges.” He has also developed a program for graphically-oriented data entry of neumed transcriptions, called the *Neume Notation Editor*. 
PHASE ONE GOALS & PROCEDURES

Initiatives. Phase One of the Project will develop the software infrastructure in several initiatives.

1. Construct a comprehensive taxonomy of neume forms including all notation genres. Developing a comprehensive taxonomy of neumes is no simple matter, but it will provide the basis for extensive comparative research in the future. The classification procedure will proceed as follows.
   1.a. List all major neume notation genres according to accepted theory in medieval musicology.
   1.b. For each genre, list all known neume forms as conceptual abstractions; correlate symbols across genres by abstract neume; construct a classification hierarchy; assign character codes.
   1.c. Find a coding solution for the problems of unheighted neumes and certainty factors.
   1.d. Construct a complete roster of auxiliary symbols used in neumed manuscripts; assign codes.
   1.e. Study the Digital Scriptorium [10] DTD tags for text manuscript transcription, and list any additional tags needed for neumed manuscripts.

2. Design and develop a data representation (including character code points and a context-free grammar) in the Private Use Area of the Unicode™ Standard to accommodate the special requirements of neume notation. The main requirements of the data representation are as follows.
   2.a. It must be comprehensive enough to support all expected uses of the data, including print and analysis. In particular, different notational genres must retain their individual characteristics in transcription, and yet it must be easy to compare melodies across genres.
   2.b. It must be designed for computational efficiency so that analytical research methods can be executed in reasonable time if many documents must be compared in a complex manner.
   2.c. It must be sufficiently formal and abstract that the data can be read and processed by programs in any programming language.

3. Design and develop a DTD in XML (called NeumeXML) that will be the ‘carrier’ for neume character data in file transfer and storage. The principal requirements of NeumeXML are as follows.
   3.a. The DTD for NeumeXML must be backward-compatible with the DTD of the Digital Scriptorium project, so that NeumeXML transcription files can be read by Digital Scriptorium applications. Dual use will avoid duplication of effort in making transcriptions.
   3.b. The DTD must include additional markup tags beyond those provided by the Digital Scriptorium project, so as to accommodate the special requirements of chant transcription.
   3.c. The DTD must be verifiable by standard XML parsers.

4. Create programs or ‘classes’ (collectively referred to as “the Programs”) that will allow any program to read, write, and modify neume character data and NeumeXML tags. Requirements are as follows.
   4.a. They shall be generalized so that they can be used in Barton’s Editor or any other program.
   4.b. They shall be written in the platform-independent Java™ language for use on any computer.
   4.c. They shall be designed and written in a strictly object-oriented manner to maximize the ease of debugging, adding of new features, and reuse in programs written by third parties.
   4.d. They shall be well documented, including design specifications and instructions for use.
   4.e. A data structure (specifically, a multi-dimensional, doubly-linked list) shall be provided for internal storage of transcription data, plus algorithms and programmed functions for list traversal, data read/insert/delete/modify, import NeumeXML, and export NeumeXML.
   4.f. A program shall be created for entry of NeumeXML meta-data such as transcriber, date, etc.
   4.g. Functions shall be provided for high-level NeumeXML operations such as get file, send file, search for neume, next verse, next syllable, next neume, read note, and compare two notes.

5. Evaluate the Programs toward the end of Phase One. Members of the Advisory Board will test the Programs using Louis Barton’s Neume Notation Editor, the Oxford Chant Group’s analysis program, or other means. The character data representation and DTD will be tested with sample transcriptions from various notational genres. Conclusions drawn from these evaluations will be included in the publication of Phase One results. During development of the Programs, Prof. Gagliardi and/or Dr. Jeavons will be totally familiar with all source code of the Programs and monitor its progress.
**Editor Program.** The *Neume Notation Editor* is a graphical, interactive Java™ Applet (which is runnable in Web browsers) for input of chant transcriptions. Louis Barton will continue developing the *Neume Notation Editor* on his own funds and time. Development of the *Neume Notation Editor* shall not be included in this phase of the Project.

**Outcomes.** Specific outcomes that the Project shall generate include the following:
- neume character definitions in the Private Use Area of the Unicode™ Standard;
- an XML document type definition (DTD) for use in file transfer and storage of neume character data;
- reusable program code that will allow any program to read, write, and modify neume character data or markup tags of the DTD;
- an informational Project Web site;
- a test set of sample transcriptions drawn from a variety of neume notational genres;
- a rudimentary database, hosted on a Harvard server, that will provide Web access to transcriptions; and
- publication of results.

**Maintenance.** On-going mediation and maintenance of the neume encoding standard and the DTD shall be by members of the Advisory Board or the Internet Assigned Numbers Authority (IANA), as deemed appropriate by the Advisory Board in consultation with other concerned parties. Louis Barton shall have on-going responsibility for maintenance of the Programs, Web site and transcription database.

**FUTURE DIRECTIONS**

**Phase Two.** Undoubtedly, field usage will reveal that improvements are needed to the software infrastructure developed under Phase One; these will be a major focus of Phase Two of the Project. An addition known to be needed is to expand the data representation to cover Byzantine chant [11].

Phase Two will explore the feasibility of inserting links in existing indexes for chant, such as the CANTUS database [5], for connecting to transcriptions. The Digital Scriptorium project [10] is writing DTDs for descriptive catalogs and text transcriptions of medieval and Renaissance manuscripts. These include hyperlink tags, which Phase Two will exploit for referencing neumed transcriptions. Also of interest is to inter-link online, digital images of manuscripts [9][17] with corresponding transcriptions.

A third goal of Phase Two will be to improve the Harvard transcription database so as to allow scholars to upload new transcriptions to it. This functionality will include user authentication, automatic validation of submitted files against the DTD, and automated indexing of accepted files in the database for search and retrieval via the Web. To develop the concept of a ‘distributed database’, whereby many institutions can put transcriptions on the Web using their own servers, the design and program code for the Harvard transcription database will be made available to other institutions. Search for a transcription on the Web will be by ordinary search engines, chant databases, or hyperlinks from related Web material.

**Phase Three.** Phase Three of the Project will involve research into optical character recognition (OCR) for rough-draft transcription from digitized manuscript images. This work will help establish the minimum requirements for high-resolution, color photography of manuscripts that will be sufficient for OCR. It is expected that some amount of operator interaction would be necessary, such as to specify the notational genre. Since chant texts recur with great stability across manuscripts, and textual calligraphy can be difficult to read, texts would likely be copied from existing transcriptions rather than optically recognized. Once the OCR program had extracted a draft transcription, a qualified person would then check it for correctness. Prof. Barton is currently consulting with DIAMM [9] about isolating neumes from background ‘noise’ in images; he expects to complete an OCR feasibility study by the time of Phase Three. Artificial intelligence techniques will likely be used for OCR of neumatic symbols in images, and for translating from ‘alphabetic’ symbols to ‘abstract’ neumes in the data representation. OCR for neumes would be enormously helpful, due to the great number of neumed documents that need to be transcribed.
MANAGEMENT PLAN

This is a unique opportunity to encourage interdisciplinary research between the humanities and sciences. An Advisory Board consisting of prominent medieval musicologists and computer scientists will participate in all major project decisions, establish evaluation criteria, and review progress. There will be two plenary meetings of the members of the Advisory Board. Further interactions will be via e-mail. The P.I. will be responsible for setting timetables, coordinating activities, fiscal management, and liaison with the Mellon Foundation. The Software Engineer will be responsible for software design and programming. He and his assistant(s) will work mainly in Missouri on his computer equipment (in the Internet age this is entirely feasible and ordinary). He will upload project materials to the Harvard server, stay in frequent e-mail contact with the P.I., and make five 3-day trips to Harvard for consultation. He will accept only part-time teaching assignments during the funded period.

Modern techniques for software project management will be used, including requirements specification, design abstraction, and software validation. The software architecture will be carefully documented. The Programs will be written in the object-oriented Java™ programming language, which (unlike C and other, older programming languages) is easy to read and maintain. This will ensure ‘transparency’ to other programmers, should it become necessary to use backup programmers.

Significant target dates for Phase One are as follows (specific dates to be announced).

- October 2001: Two-day meeting of the Advisory Board at Harvard. Agenda: complete a taxonomy of neumes; discuss problems and strategies for Phase One; discuss the management plan and the role of Advisory Board members in decision-making; draw up guidelines for progress review.
- March 2002: Rollout of alpha version of neume data representation in the Private Use Area of the Unicode™ Standard. Rollout of an alpha version of the Java™ programs for evaluation and testing by the Advisory Board members. Rollout of a Web site as central repository for Project information.
- August 2002: Present the alpha data representation at the biennial symposium of the Cantus Planus Study Group of the International Musicological Society, to be held in Leuven, Belgium.
- September 2002: Rollout of alpha version of the NeumeXML DTD. Rollout of Java™ program additions for reading and writing of NeumeXML files and meta-data.
- October 2002: Second meeting of the Advisory Board at Harvard. Agenda: present progress report; draft the Board’s evaluation report; discuss plan for future work and funding.

BUDGET

[omitted from the online version]
Appendix A: References Cited

[15] Early Manuscripts at Oxford University, at *http://image.ox.ac.uk/.
[34] ————, *Paléographie musicale*, (Sablé-sur-Sarthe, France: Abbaye Saint-Pierre de Solesmes, 1889-).

Digital Encoding for Medieval Chant Transcription; proposal, July 17, 2001 9
Appendix B: Curricula Vitae of Principal Personnel

Prof. Thomas Forrest Kelly (Principal Investigator)
Professor of Music and Chair of the Department of Music, Harvard University
e-mail: tkelly@fas.harvard.edu

EDUCATION

EMPLOYMENT
1999- Chair, Department of Music, Harvard University
1994- Professor of Music, Harvard University
1988-94 Oberlin College Conservatory of Music: Professor and Dana Faculty Fellow; Chair, Historical Performance Program; Acting Dean of the Conservatory, 1990-91.
1982-84 Amherst College: Associate Professor of Music and Director of the Five College Early Music Program.
1979-82 Smith College: Associate Professor of Music and Director of the Five College Early Music Program, 1981-1982; Assistant Professor of Music and Director of the Five College Early Music Program, 1979-1981.

AWARDS
Otto Kinkeldey Award of the American Musicological Society for the most distinguished work of musicological scholarship of 1989.

SELECTED PUBLICATIONS
Early Music in America, (Early Music America, 1989).

RECENT COLLABORATORS
The Monumenta Liturgical Beneventana project at the Pontifical Institute of Mediaeval Studies, University of Toronto.
The Centre Européen pour la Recherche et l’Interprétation des Musiques Médiévales, Fondation Royaumont, France.
Louis W. G. Barton (Software Engineer)
University of Oxford (external); contract programmer
e-mail: louis.barton.je.77@aya.yale.edu

EDUCATION
Harvard University, S.M. in Computer Science, 1998.
Webster University (St. Louis), B.S. in Computer Science, 1996, with departmental honors.
Yale University, B.A. in Music and Computer Science, 1995, cum laude and with distinction in the major.

EMPLOYMENT
2000 summer Lecturer in Computer Science, Harvard University Summer School.
1999 - 2000 Assistant Professor of Mathematics and Computer Science, Suffolk University (Boston).
1998 - 1999 Visiting Assistant Professor of Mathematics and Computer Science, Suffolk University.
1995 - 1996 Instructor in Computer Information Systems, National College (Kansas City, MO).
1980 - 1993 Programmer/Analyst, Cecilia House, Inc. (Theodosia, MO). Contract programming for Impact Hearing Conservation, Inc. (Kansas City, MO; 4 years); Micro Apothecare, Inc. (Gassville, AR; 5 years); Claritas Partners, Inc. (Alexandria, VA; 3 years). Database design and user-interface programming in C and Pascal.
1968 summer Programmer Trainee, IBM Corp. (Harrison, NY). Data-analysis functions in PL-1.

PUBLICATIONS
“Ternary certainty factors for transcription data streams,” (forthcoming).

PROFESSIONAL AFFILIATIONS
Association for Computing Machinery (ACM), member.
Association for Computers and the Humanities (ACH), member.

OTHER
Prof. Barton has been interested in computer applications to music since 1967. In 1972 he was the first to design and build an interface between an electronic organ and a computer (long before MIDI was invented for this purpose). At Yale, his Senior Project involved adapting a commercial music-editing program to print chant notation using a character font that he designed. At Harvard, his master’s degree ‘thesis’ involved research on the feasibility of a lossless data representation for medieval neume notation. While a student at M.I.T, he did a proof-of-concept optical character recognition program for handwritten neumes. His doctoral thesis at Oxford is on coding-design theory for nondeterminate symbol systems.
Appendix C: Samples from Neumed Documents  (showing evolution of neumatic writing)

Sample #1: Beneventan notational genre, no line. Detail of Benevento, Biblioteca capitolare, MS 38, f. 53. [Reproduced from Kelly, *The Beneventan Chant*, pl. 18.]

Sample #2: Messine notational genre, no line; ca. early 10th century. Laon, Bibliothèque Municipale 239. [Reproduced from Solesmes, *The Musical Notation of Latin Liturgical Chants*, pl. 17, p. 34.]
Sample #3: Early Italian (Bologna) notational genre, no line; second half of 11th century. Rome, Biblioteca Angelica, MS 123, fol. 126 (26.5x17.5 cm). [Reproduced from Stäblein, *Schriftbild der Einstimmigen Musik*, pl. 17b, p. 127.]
Sample #4: Beneventan notational genre, single line. Benevento, Biblioteca capitolare, MS 21 f. 235v. [Reproduced from Kelly, *The Benevantan Chant*, pl. 11.]
Sample #5: German Hufnagel notational genre, lined; 15\textsuperscript{th} century. Leaf from a manuscript Bible. [Reproduced from Beverly A. Barksdale, \textit{The Printed Note}, (Toledo, OH: Toledo Museum of Art, 1957), No. 3, p. 6.]
Sample #6: Square-neume notational genre, four lines; A.D. 1440. Detail from Metten (Lower Bavaria), Bibliothek der Benediktinerabtei, fol. 64v. [Reproduced from Stäblein, *Schriftbild der Einstimmigen Musik*, pl. 45a, p. 164.]

Sample #7: Stencil-painted, giant square notation, four lines; A.D. 1732. Solesmes, MS 34 p. 56. [Reproduced from Solesmes, *The Musical Notation of Latin Liturgical Chants*, pl. 29, p. 49.]
Appendix D: Glossary of Musicological Terms

antiphoner  A book, usually manuscript, containing music for use at the Mass (antiphonale missarum) or the daily office (antiphonale officii).

clivis     A two-note neume in which the second note is lower than the first.

diatonic  A musical scale (as is used in Western music) where notes fall into discrete pitches that are a whole tone or a semitone apart.

height    The vertical placement of a neume symbol above the Latin text. It is convenient to think of 'height' roughly in terms of note pitch. Especially in early neumed documents, however, neume height might only be a general indicator of melodic direction.

illumination Artistic decoration of manuscripts that is not an essential part of the text or melody. The category includes florid initial letters, pictorial ‘miniatures’, border motifs, and other decorations.

liturgy   The body of religious rituals prescribed for community worship. When referring to the Liturgy of the Catholic Church, it is comprised mainly of the Mass and the Canonical Hours.

mensural notation Musical notation that prescribes specific relative durations of notes. In modern notation this is done with quarter notes, half notes, etc., assembled into units called 'measures'. Medieval chant notation does not specify such rhythmic values for notes.

neume One of the set of symbols (simple or compound) used in medieval times for writing music. Neumes are quite different from modern musical symbols, both in appearance and meaning.

oriscus A neume modifier or a free-standing neume form, written in some shape similar to the tilde of some modern languages, to indicate a variety of nuances in the vocal rendering.

pictographic notation Musical notation depicting general melodic shape, rather than specific notes.

pressus A compound neume consisting of a virga, an oriscus, and a punctum. The pressus can be found either in isolation or in combination.

quilisma Two or three semicircular loops, never found in isolation but is always tied to an ascending virga forming a quilisma-pes. Normally a quilisma-pes is preceded by a note which is always at a lower pitch. Most often this note is part of a quilismatic group but sometimes it belongs to the preceding neume. In this case the quilisma is directly attached to a new text syllable [6].

staff     One or more horizontal lines on which music notation is written, where higher lines indicate higher pitch. In medieval manuscripts, however, there might be no staff line, or between one and six lines per staff. Modern music uses five staff lines.

virga    A single-note neume written in a manner similar to an acute accent.