

# **21M.269 – Studies in Western Music History:**

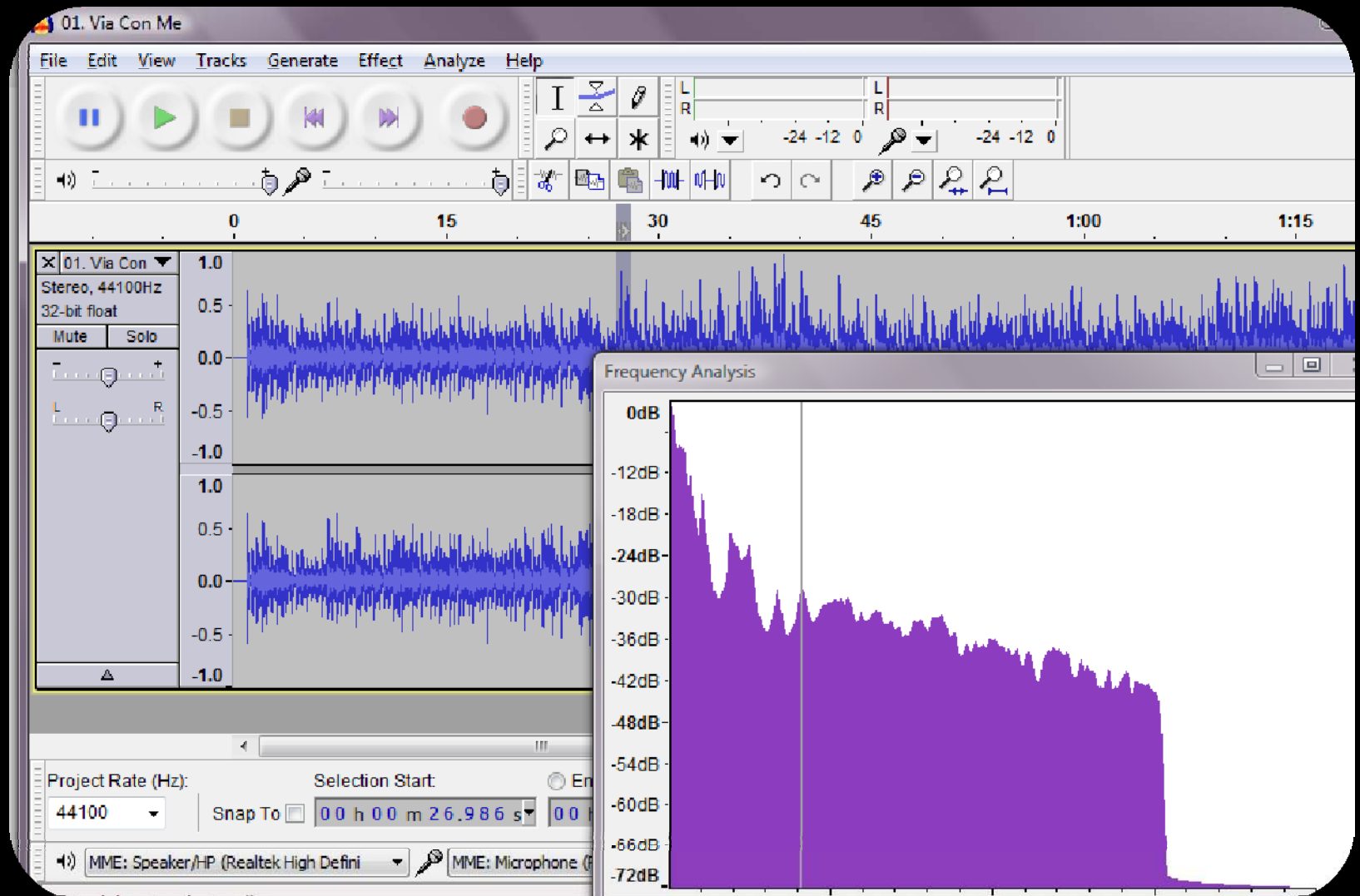
## **Quantitative and Computational Approaches to Music History**

Michael Scott Cuthbert  
Associate Professor of Music, M.I.T.

*7 February 2012*

[Screenshot of iTunes® removed due to copyright restrictions.]

Computers are changing how we acquire and share music



Using computers to analyze and manipulate audio

20100 - [residency\_week\_in2010.mus]

File Edit Utilities View Document MIDI/Audio Plug-ins Tools Window Help

43

bas clarinet to cl.

low tom ped bd toms tom rim only w.b. temple blocks 7 cowb.

hit Eb then bow Ab

vibes.

E. Gu.

Pf.

Vc.

Cb.

Vasarely Patterns, p. 3

Finale® user interface © MakeMusic, Inc. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <http://ocw.mit.edu/fairuse>.

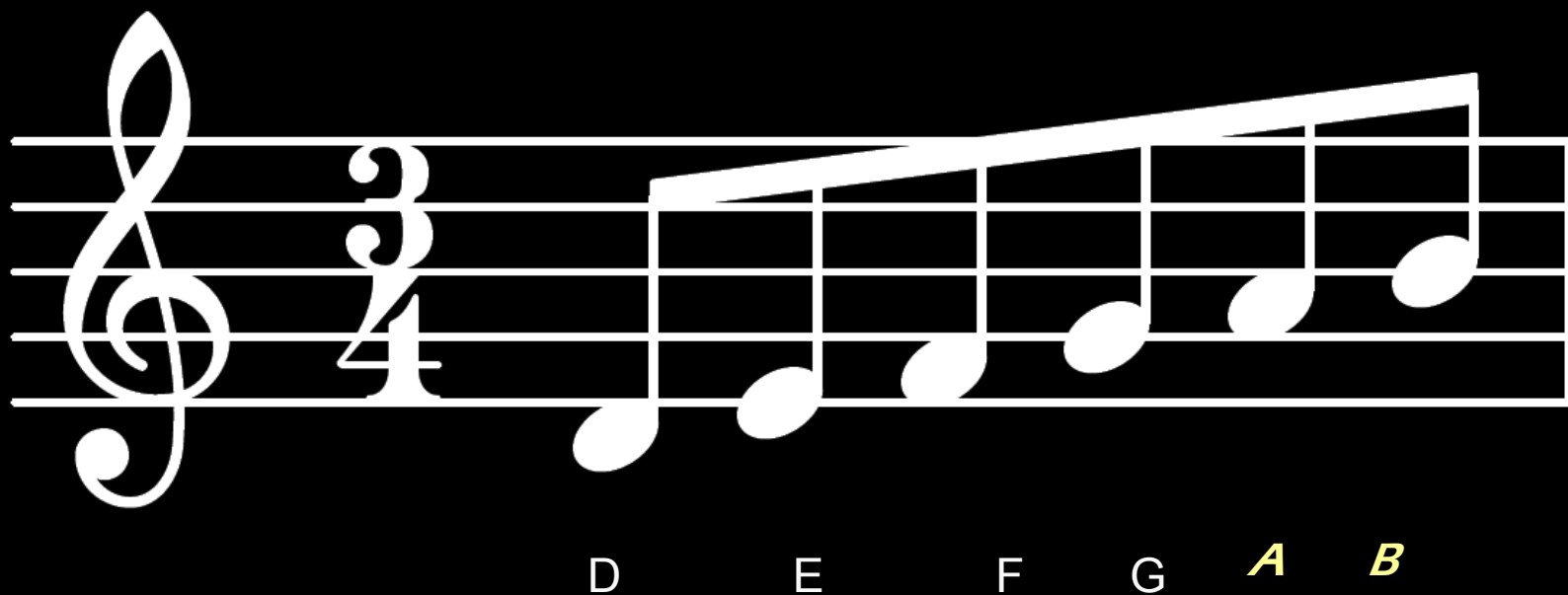
Computers are changing how composers notate their scores

# Despite all the advances in using Computers in Music

practically nothing has changed in  
how computers are used to analyze scores.

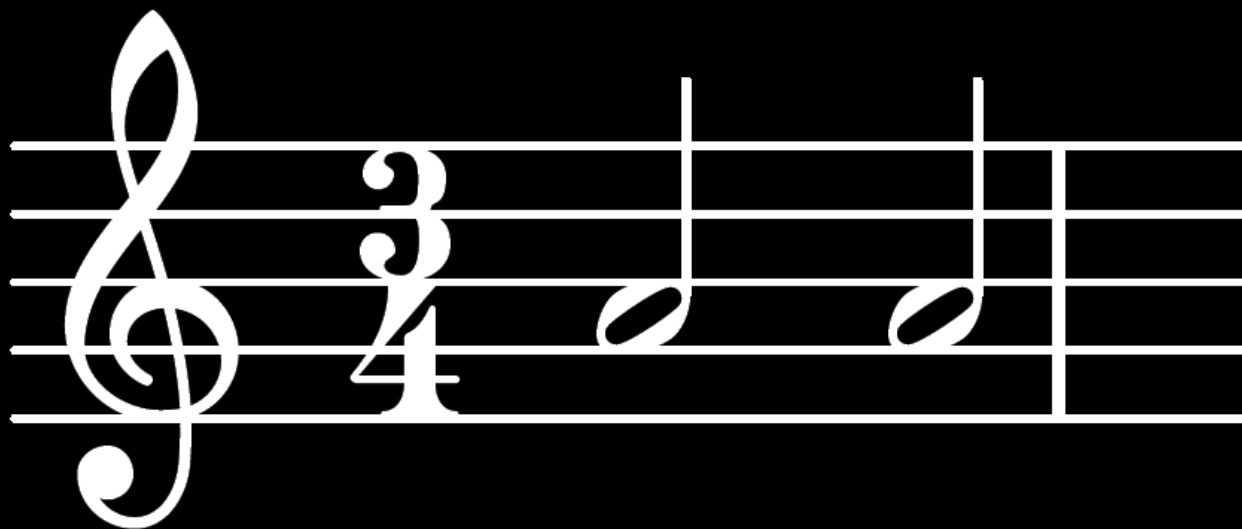
# Simple things that a computer should know about music but doesn't...

Scales ...



# Simple things that a computer should know about music but doesn't...

Meter...



# Simple things that a computer should know about music but doesn't

Elementary Harmony:  
Which of these is more dissonant?

Stravinsky *Rite of Spring*

The image shows a musical score for three staves from Stravinsky's *Rite of Spring*. Each staff begins with the instruction "arco (non div.)" and "tutti". The first staff has a dynamic marking of "f". The score is divided into three measures. The first measure contains a series of notes. The second measure is marked "sempre stacc." and contains notes with stems pointing downwards. The third measure is marked "sempre simile" and contains notes with stems pointing upwards. The overall dynamic is "f".

The image shows a musical score for the piano part of Mozart's *Piano Sonata in C major*. It features a treble and bass clef. The treble clef part starts with a dynamic marking of "mp" and a fingering of "1 3 5". The bass clef part starts with a dynamic marking of "p" and a fingering of "1 2". The score is divided into three measures. The first measure contains a series of notes. The second measure is marked "p" and contains notes with stems pointing downwards. The third measure is marked "p" and contains notes with stems pointing upwards. The overall dynamic is "p".

Mozart *Piano Sonata in C major*



# 1812 Ouverture solennelle

Arranged by Stepán Esipoff

P. Tschaiḱowsky

Largo.

The image shows the first system of a musical score for the 1812 Overture. It consists of two staves: a treble clef staff (piano) and a bass clef staff (bass). The time signature is 3/4. The key signature has two flats (B-flat and E-flat). The tempo is marked 'Largo.' and the dynamics include 'mf' (mezzo-forte) and 'f' (forte). There are several 'ten.' (tension) markings above and below the staves. The score includes various musical notations such as slurs, ties, and asterisks. The bass staff has some markings that look like 'Led.' and 'sfz' (sforzando). The overall style is a grand staff with a focus on harmonic texture and dynamics.

The problem with doing large-scale analysis is that in order to say something useful about, say, the key of a piece, you need a system that can identify and work with a phrase...

1812

# Overture solennelle

Arranged by Stepán B...

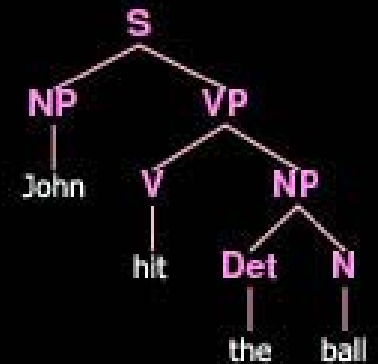
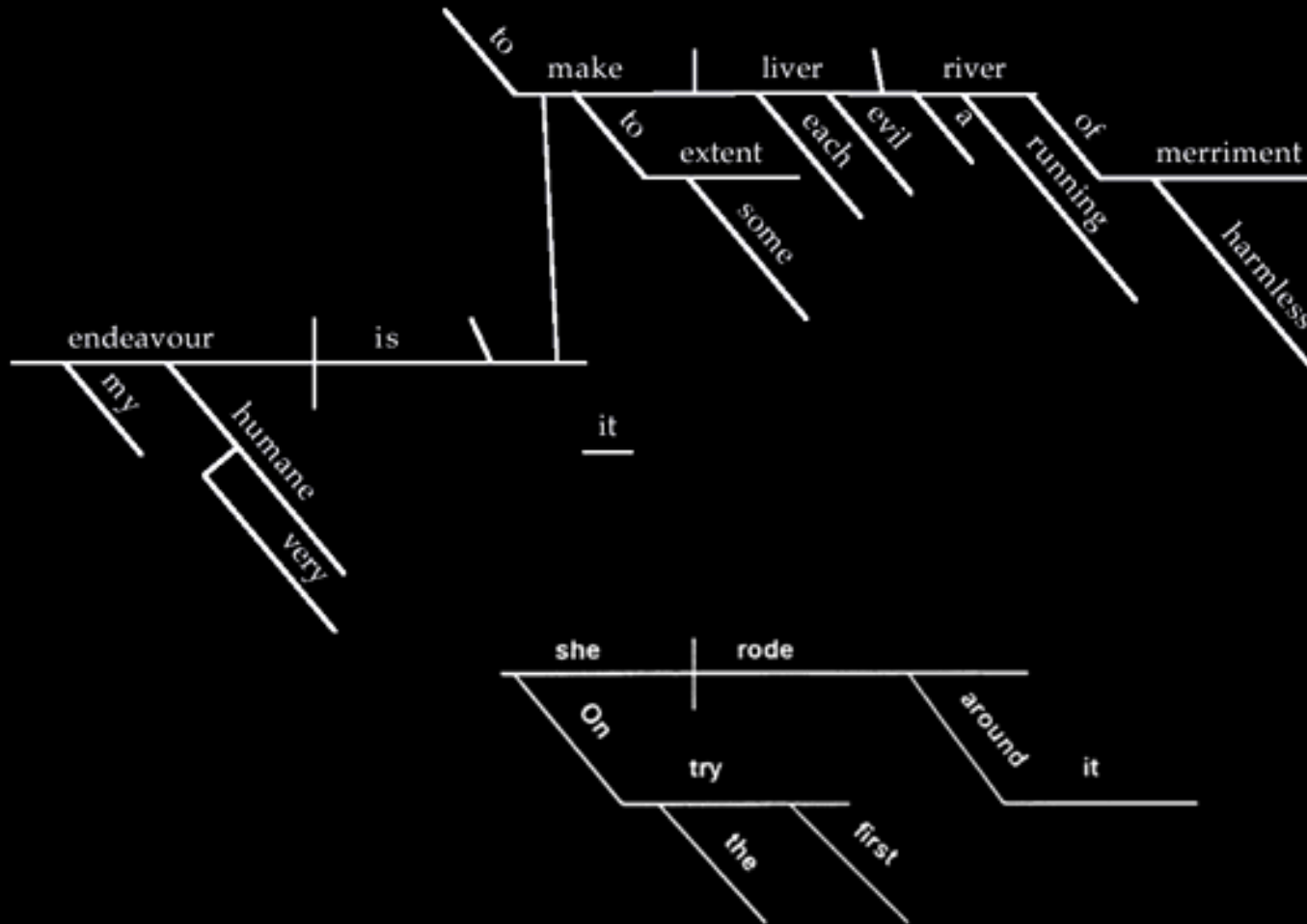
Larg

P. Tschaïkowsky

A page of musical notation for the Overture solennelle by P. Tschaïkowsky. The score is arranged for piano and features two systems of music. The first system includes a treble and bass clef staff with various notes, rests, and dynamic markings such as 'ten.' and 'f'. A grey rectangular box highlights a specific section of the first system. The second system continues the musical piece with similar notation and dynamic markings like 'cresc.', 'ff', and 'sfz'. The page is annotated with several asterisks and 'ten.' markings, likely indicating specific points of interest or technical challenges.

...and in order to work with phrases you need an object model that understands chord progressions. And in order to work with progressions you need to be able to classify individual chords...and to do that you need a robust way of modeling notes. So a small problem becomes a bigger and bigger problem.

# Hypothetical Example from Computer-aided Linguistics: Using Sentence Structure to find Spam E-Mails



# Hypothetical Example from Computer-aided Linguistics: Using Sentence Structure to find Spam E-Mails

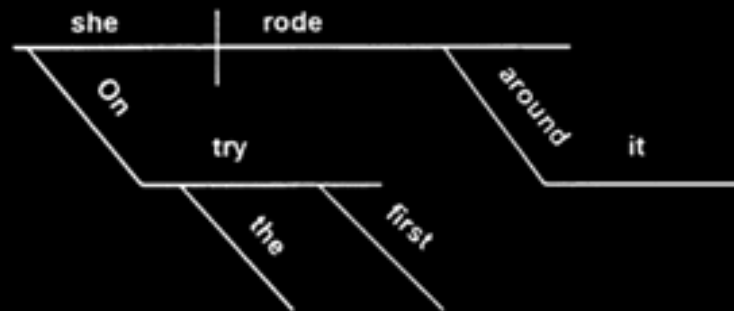
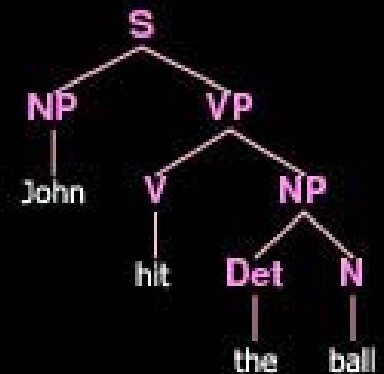
*Available Packages for Sentence Parsing (a quick search):*

Stanford Natural Language Processing Group

Apple Pie Parser (NYU)

SenDraw @ University of Central Florida

OpenNLP sentence parser



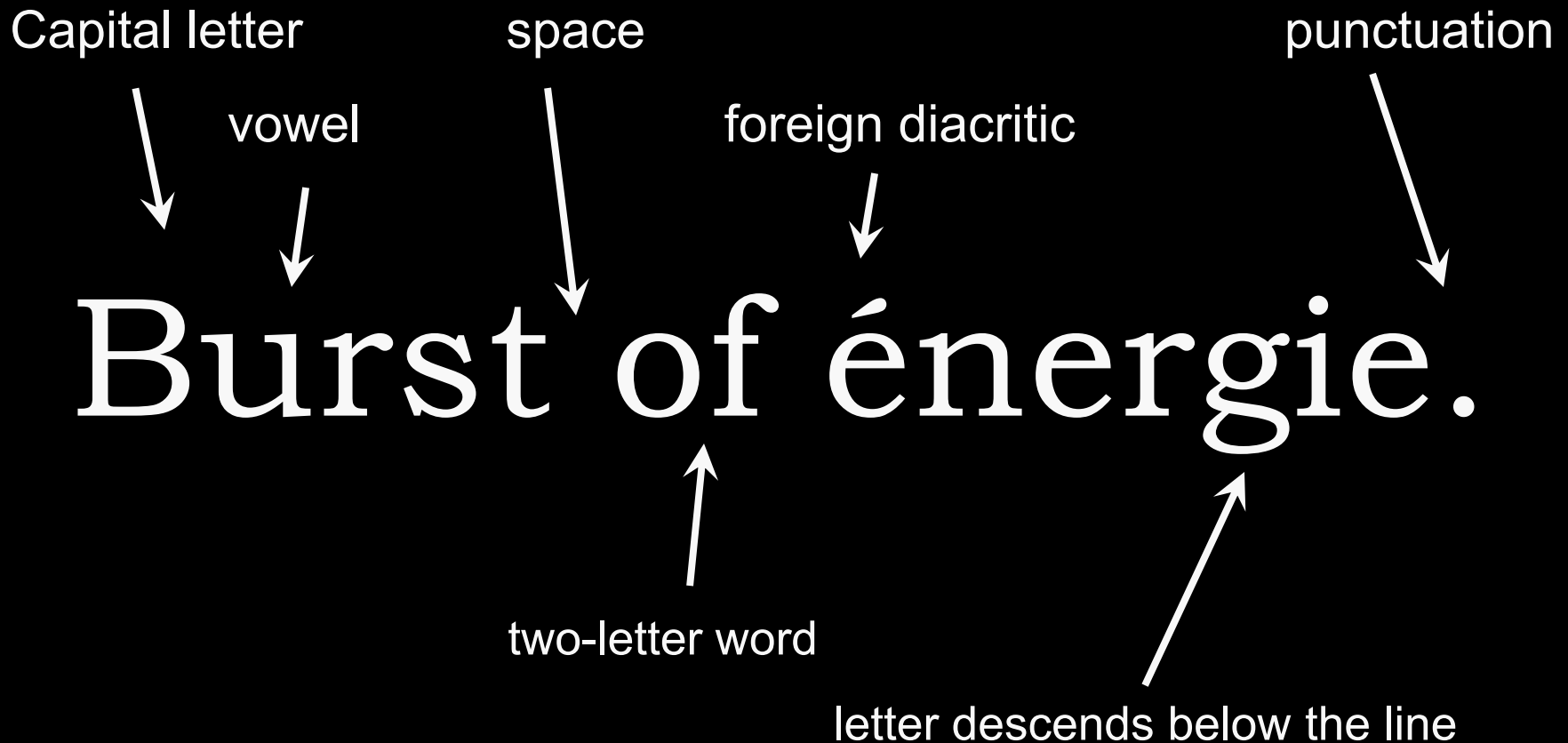
# Hypothetical Example from Computer-aided Linguistics: Using Sentence Structure to find Spam E-Mails

[Screenshots of online thesarus, dictionary, and word lists removed due to copyright restrictions.]

Hypothetical Example from Computer-aided Linguistics:  
Recognition of text characteristics that we take for  
granted

Burst of énergie.

Hypothetical Example from Computer-aided Linguistics:  
Recognition of text characteristics that we take for granted



Examples of Humdrum (David Huron) code:

**Problem:** “Do the words ‘high’, ‘hoch’, or ‘haut’ coincide with high notes in the text?”

**Solution:** Compare “semits \* | text | extract -i '\*\*semits,\*\*text' | ditto -s = \ |  
egrep -i '^\*|high|hoch|haut' | extract -i '\*\*semits' | stats” to “semits \* |  
extract -i '\*\*semits' | ditto -s = | rid -GLI \ | stats”

**Problem:** “Convert one of Arnold Schoenberg’s 12-tone rows to a 12-tone matrix.”

**Solution:** X=0 ; while [ \$X -ne 12 ] ; do ; reihe -a -P \$X \$3 | rid -GLId | head  
-\$2 > P\$X ; reihe -a -I \$X \$3 | rid -GLId | head -\$2 > I\$X ; reihe -a -R \$X \$3  
| rid -GLId | head -\$2 > R\$X ; reihe -a -RI \$X \$3 | rid -GLId | head -\$2 >  
RI\$X ; let X=\$X+1 ; done



# music21: a toolkit for computer-aided musicology

## What is music21?

Music21 is a set of tools for helping scholars and other active listeners answer questions about music quickly and simply. If you've ever asked yourself a question like, "I wonder how often Bach does *that*" or "I wish I knew which band was the first to use these chords in this order," or "I'll bet we'd know more about Renaissance counterpoint (or Indian ragas or post-tonal pitch structures or the form of minuetts) if I could write a program to automatically write more of them," then music21 can help you with your work.

## How simple is music21 to use?

Extremely. After starting [Python](#) and typing "from music21 import \*" you can do all of these things with only a single line of music21 code:

### Display a short melody in musical notation:

```
tinyNotation.TinyNotationStream("c4 d8 f g16 a g f#", "3/4").show()
```

### Hear the interval of a quarter tone (coming soon!):

```
interval.ChromaticInterval(0.5).show("midi")
```

### Print the twelve-tone matrix for a tone row (in this case the opening of Schoenberg's Fourth String Quartet):

```
print (serial.rowToMatrix([2,1,9,10,5,3,4,0,8,7,6,11]) )
```

or since all the 2nd-Viennese school rows are already available as objects, you can type:

```
print (serial.RowSchoenbergOp37().matrix() )
```

### Convert a file from Humdrum's \*\*kern data format to MusicXML for editing in Finale or Sibelius:

```
parse('/users/cuthbert/documents/composition.krn').write('xml')
```

With five lines of music21 code or less, you can:

```
def closedPosition(self):
    ...
    returns a new Chord object with ...

>>> chord1 = Chord(["C#4", "G5",
>>> chord2 = chord1.closedPosition
>>> print(chord2.lily.value)
<cis' e' g'>4
...

newChord = copy.deepcopy(self)
tempChordNotes = newChord.pitches
chordBassPS = self.bass().ps
for thisPitch in tempChordNotes:
    while thisPitch.ps > chordBassPS:
        thisPitch.octave = thisPitch.ps + 1
newChord.pitches = tempChordNotes
```

- [Get Started](#) with music21
- [Browse the music21 documentation](#)
- [Download music21](#) from Google Code
- [Get our latest news](#) and updates at the music21 blog
- Read the [Frequently Asked Questions](#) list
- [Sign up](#) for the music21list mailing list through Google Groups.

LATESTMUSIC21 NEWS:

### music21 at ISMIR

Chris and I will be presenting music21 as an oral paper and poster at the [International Society for Music Information Retrieval](#) in Utrecht this August. We hope we're able to meet many of you there! Chris will also be talking about the meterical features of the system at the [International Computer Music Conference](#) in NY in early

# music21: libraries for importing and exporting many formats

*In only*

abc

MuseData Stage 2

MD Stage 1

Humdrum/KERN

NoteworthyComposer

romanText (Tymoczko or Clercq/Temperley flavors)

Monophonic audio

*In & Out*

MusicXML

MIDI

Scala

Braille

*Out only*

Lilypond

Orange

Weka

# Simpler

Problem: “Convert one of Arnold Schoenberg’s 12-tone rows to a 12-tone matrix.”

```
Humdrum Solution: X=0 ; while [ $X -ne 12 ] ; do ; reihe -a -P $X  
$3 | rid -GLId | head -$2 > P$X ; reihe -a -I $X $3 | rid -GLId | head  
-$2 > I$X ; reihe -a -R $X $3 | rid -GLId | head -$2 > R$X ; reihe -a -  
RI $X $3 | rid -GLId | head -$2 > RI$X ; let X=$X+1 ; done  
(produces 48 files on disk)
```

# Simpler

Problem: “Convert one of Arnold Schoenberg’s 12-tone rows to a 12-tone matrix.”

0	2	11	7	8	3	9	1	4	10	6	5
10	0	9	5	6	1	7	11	2	8	4	3
1	3	0	8	9	4	10	2	5	11	7	6
5	7	4	0	1	8	2	6	9	3	11	10
4	6	3	11	0	7	1	5	8	2	10	9
9	11	8	4	5	0	6	10	1	7	3	2
3	5	2	10	11	6	0	4	7	1	9	8
11	1	10	6	7	2	8	0	3	9	5	4
8	10	7	3	4	11	5	9	0	6	2	1
2	4	1	9	10	5	11	3	6	0	8	7
6	8	5	1	2	9	3	7	10	4	0	11
7	9	6	2	3	10	4	8	11	5	1	0

Music21 solution:

```
print serial.rowToMatrix([0,2,11,7,8,3,9,1,4,10,6,5])
```

# Built in Analytical Libraries

## *Metrical Analysis*

```
>>> from music21 import *
>>> bwv11_6 = corpus.parse('bach/bwv11.6.xml')
>>> alto = bwv11_6.parts['alto']
>>> excerpt = alto.measures(13,20)
>>> analysis.metrical.labelBeatDepth(excerpt)
>>> excerpt.show()
```

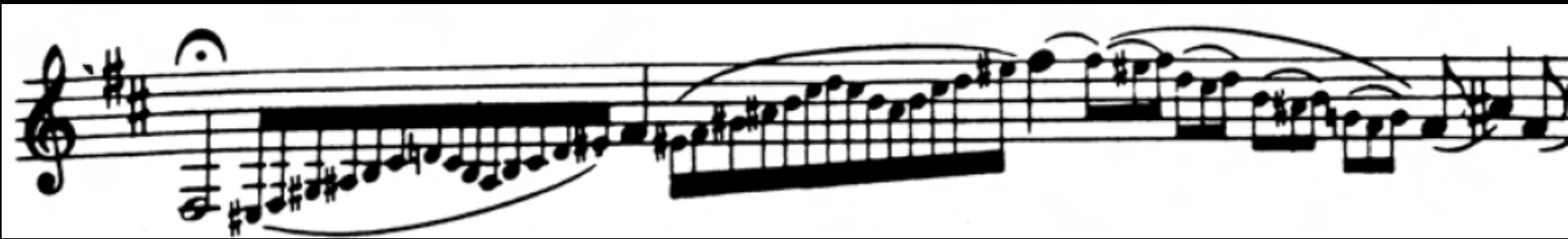
The image displays a musical score for an alto part in G major, 3/4 time, measures 13-20. The score is written on a single staff with a treble clef. The key signature is one sharp (F#) and the time signature is 3/4. The melody consists of eighth and quarter notes, with some rests. Below the staff, asterisks are placed to indicate the metrical depth of each note, showing how they align with the 3/4 beat structure. The asterisks are arranged in a grid-like pattern, with each note having a vertical column of asterisks below it, representing its depth in the measure.

# Problems of automated CMN analysis

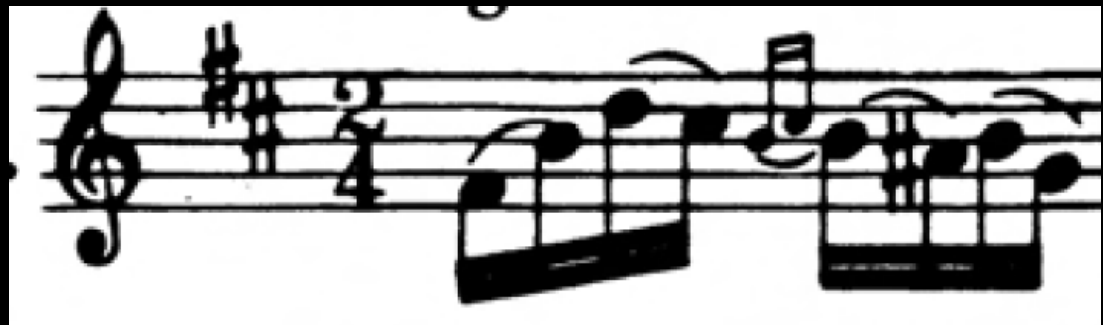
Pickups



Cadenzas



On the  
beat



Before  
the beat

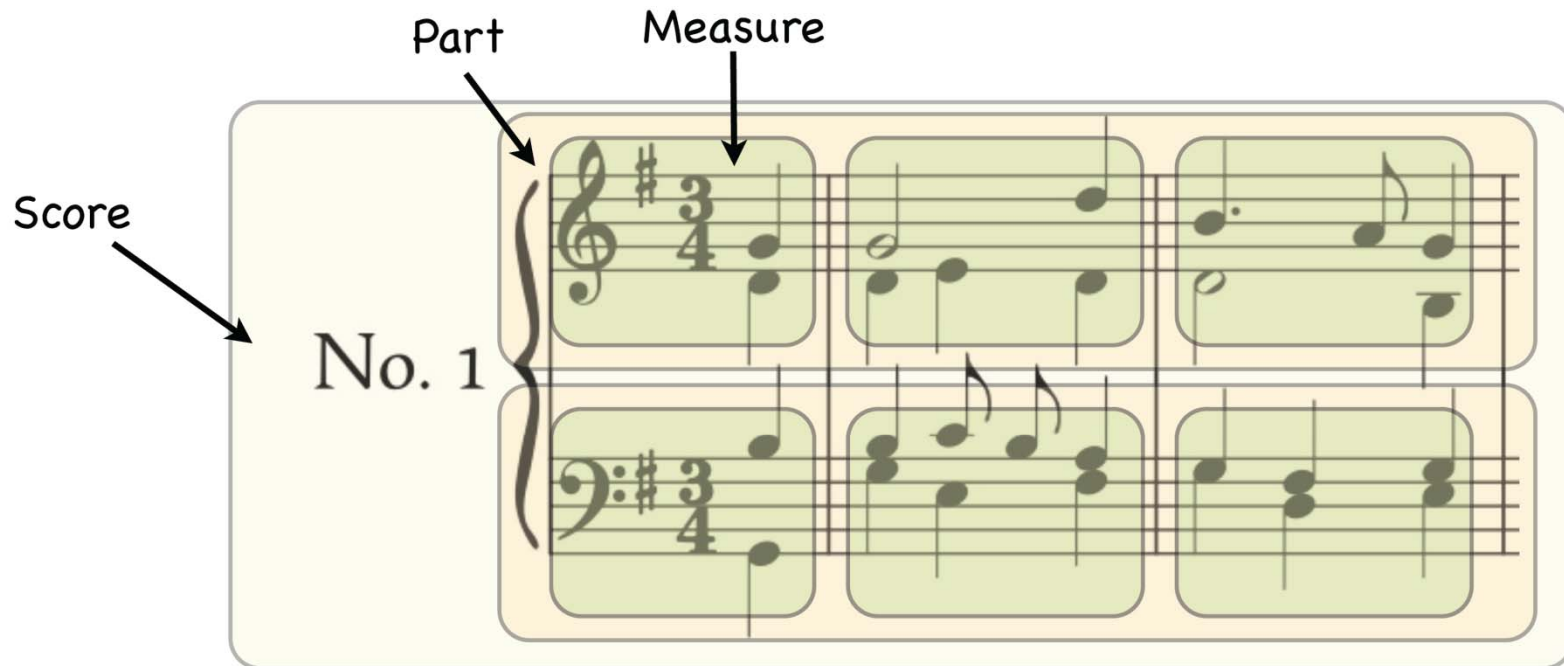


# REPRESENTATION

The next 3 slides are by William Andrew Burnson

A pervasive approach to representing music comprises part-based (row first) multidimensional containers.

Part-Based Containers (i.e. measures in parts in score)



Courtesy of William Andrew Burnson. Used with permission.



# REPRESENTATION

Alternatively, a time-based (column first) container can be used

## Time-Based Containers (i.e. instants in parts in score)

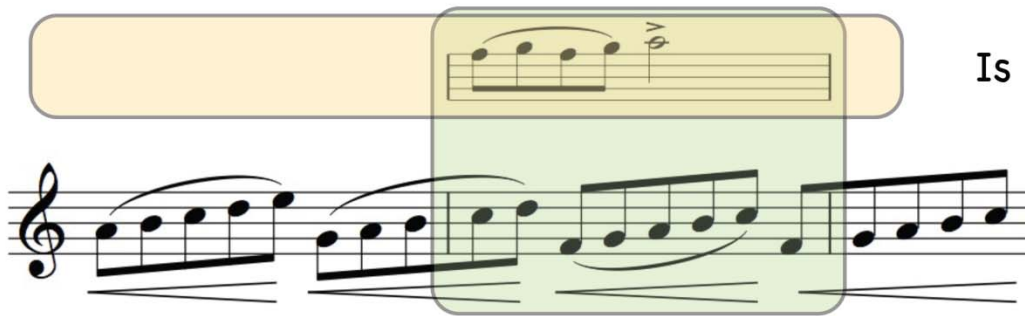
The diagram shows a musical score for two parts, labeled 'No. 1', in 3/4 time with a key signature of one sharp (F#). The score is presented in a time-based container format, where each vertical column represents an instant in time. The top staff is the treble clef and the bottom staff is the bass clef. A red dashed horizontal line separates the two staves. The score is enclosed in a light yellow rounded rectangle. Labels with arrows point to the 'Part' (treble clef), 'Instants' (a vertical column), and 'Score' (the entire score area).

Courtesy of William Andrew Burnson. Used with permission.

# REPRESENTATION

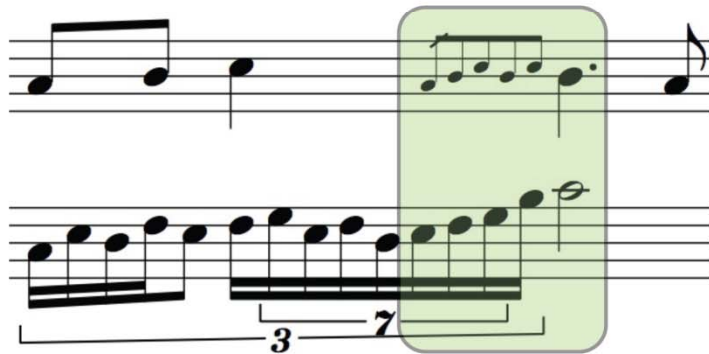
Caveats of a container-based representation:

1) Example of Part-Based failure: Ossia Staff



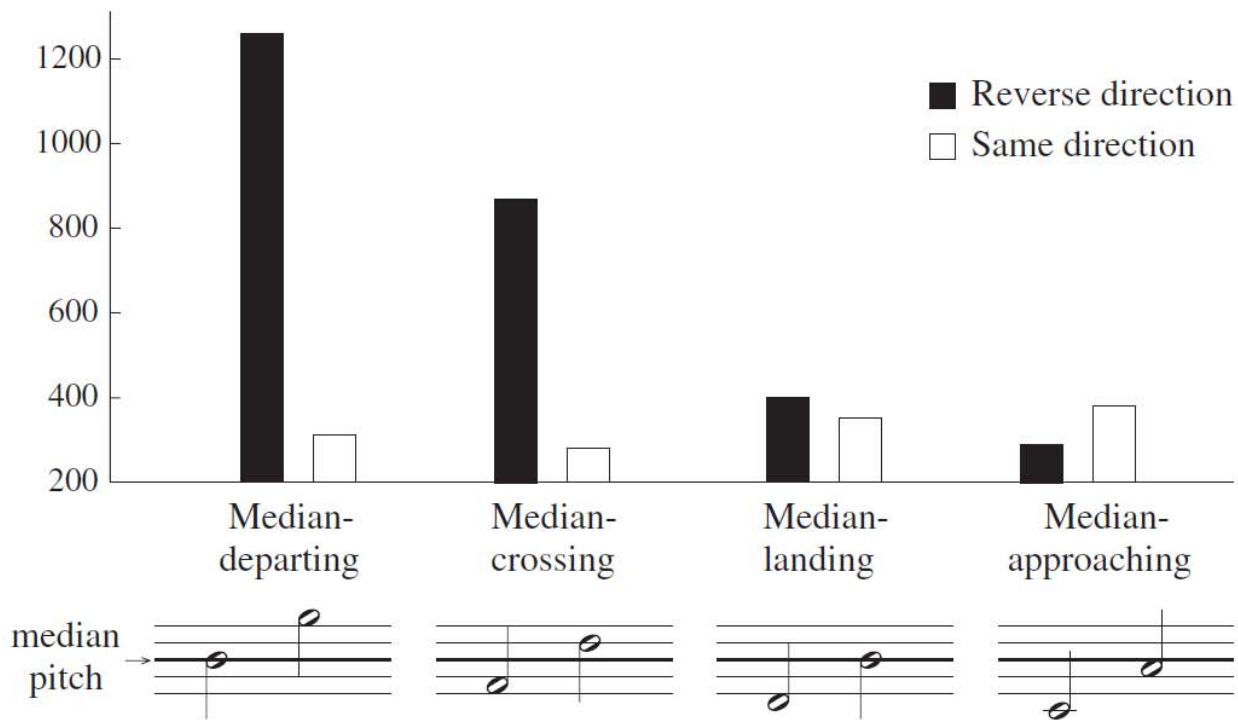
Is this a part with hidden measures?

2) Example of Time-Based failure: Grace notes



In what order should the notes appear?

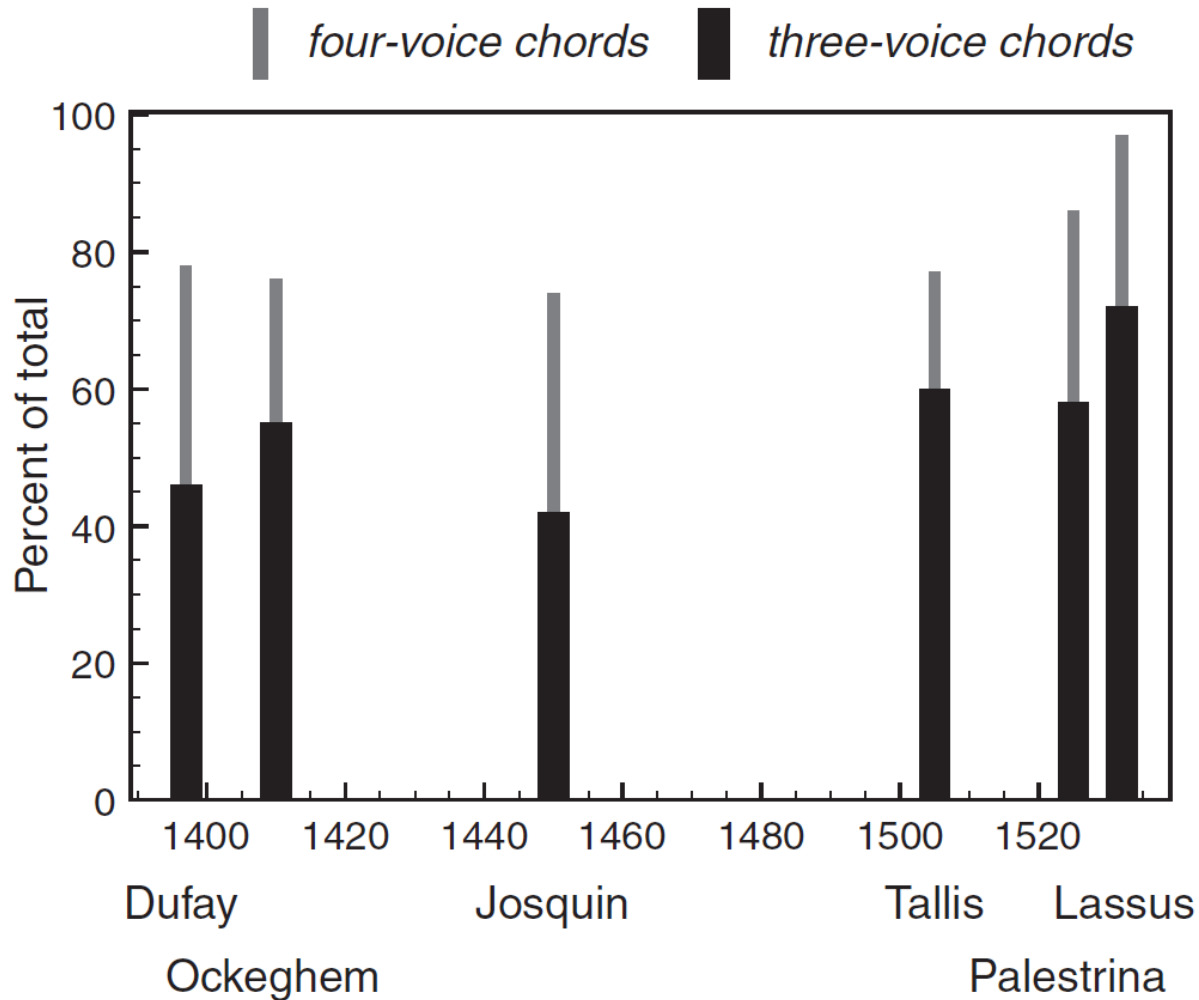
If it's such a pain, why do it?



**Figure 5.5**

Number of instances of various melodic leaps found in a cross-cultural sample of melodies. Most large intervals that approach the median pitch continue in the same melodic direction. Large intervals that land on the median pitch are as likely to continue in the same direction as to reverse direction. Results support the phenomenon of melodic regression and fail to support post-leap reversal. From von Hippel and Huron 2000.

**Figure 6.3.5** The percentage of consonant sonorities that are complete triads rather than “doubled” intervals, by composer date of birth. The percentage is reasonably high throughout the Renaissance, and increases over time.



# Some domains for Computational Musicology

Mozart accompaniment patterns

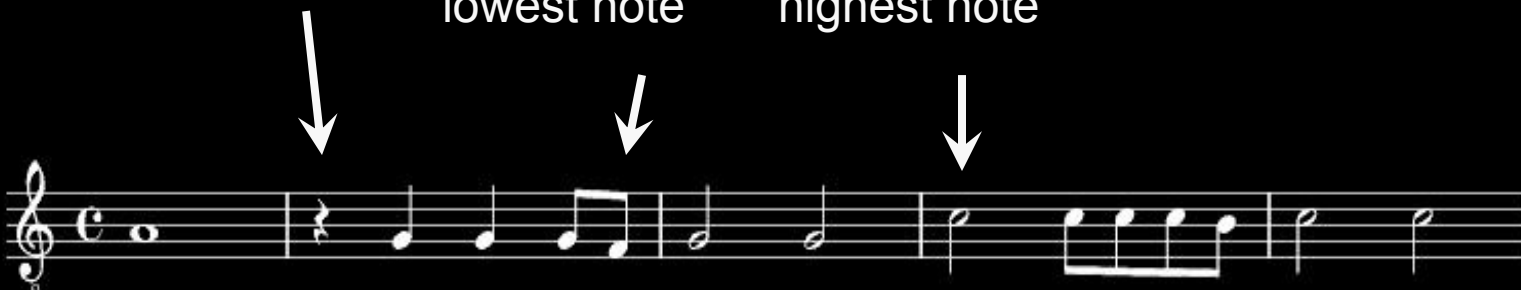
Aaron Copland's harmonies

GIS and Gregorian Chant

Authenticity and Bach's Works

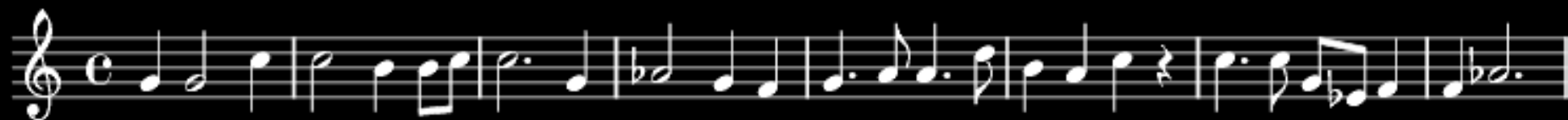
# Cavalli, *Gli amori d'Apollo e di Dafne*

Gap                      Small range  
                                 lowest note                      highest note



*Giá*                      *dell' Al – ba vi – ci - na*                      *L'au- re pre-cor-ri-tri - ci.*  
Already daybreak is near                      The morning air outruns us.

# Gluck, *Orfeo and Euridice* + Random notes



*Gluck's pitches*



*Randomly generated*



# Some domains for Computational Musicology

Mozart accompaniment patterns

Aaron Copland's harmonies

GIS and Gregorian Chant

Authenticity and Bach's Works

Baroque Opera Arias

Randomness and Musical Style

Algorithmic Composition

# Bach usually ends minor-mode pieces in major. When does he end them in minor?

```
results = stream.Stream()
for chorale in corpus.bachChorales:
    analyzedKey = chorale.analyze('key')
    if analyzedKey.mode == 'minor':
        lastChordPitches = []
        for part in chorale.parts:
            lastChordPitches.append(part.flat.pitches[-1])
        lastChord = chord.Chord(lastChordPitches)
        lastChord.duration.type = "whole"

    if lastChord.isMinorTriad() is True:
        m = stream.Measure()
        m.keySignature = chorale.flat.getElementsByClass('KeySignature')[0]
        m.append(lastChord)
        results.append(m)
results.show()
```



## Creating a Reduction and Labeling Intervals

This example, after parsing a polyphonic work stored as an Opus and creating a Score, presents and labels the intervals of the resultant chords of all distinct harmonies:

```
from music21 import corpus
# Parse an Opus, a collection of Scores
o = corpus.parseWork('josquin/laDeplorationDeLaMorteDeJohannesOckeghem')
# Create a Score from a Measure range
sExcerpt = o.mergeScores().measures(127, 134)
# Create a reduction of Chords
reduction = sExcerpt.chordify()
# Iterate over the Chords and prepare presentation
for c in reduction.flat.getElementsByClass('Chord'):
    c.closedPosition(forceOctave=4, inPlace=True)
    c.removeRedundantPitches(inPlace=True)
    c.annotateIntervals()
# Add the reduction and display the results
sExcerpt.insert(0, reduction)
sExcerpt.show()
```



- Quantity
- Searching the Corpus by Locale
- Finding Chords by Root and Collecting their Successors
- [Pitch and Duration Transformations](#)
- Basic Counting of and Searching for Musical Elements

Previous topic

Overview: Meters, Time Signatures, and Processing Beams, Accents, and Beats

Next topic

Installing Music21

This Page

Show Source

Quick search

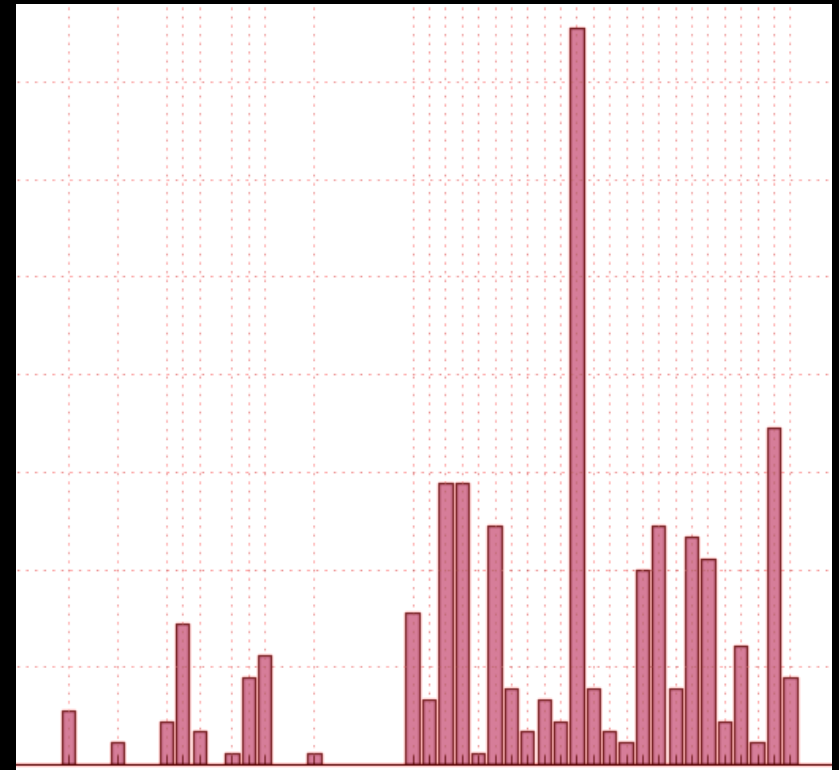
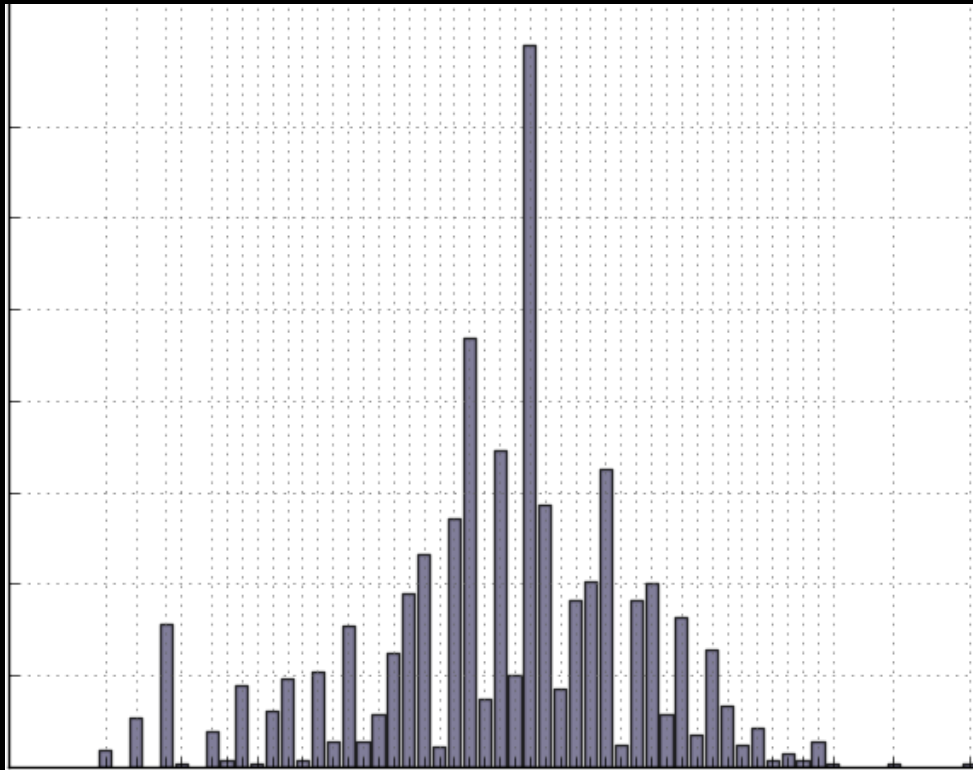
A musical score for three voices and two basses, consisting of six staves. The top three staves are for voices (Soprano, Alto, Tenor) and the bottom two are for basses. The music is written in a common time signature (C) and features a variety of note values including quarter, eighth, and half notes, as well as rests. Phrasing slurs are used to indicate melodic lines across multiple measures. The bass part includes several measures with whole rests, indicated by a small square on the staff line.

Musical score for three treble clefs and two bass clefs. The first three staves are treble clefs, and the last two are bass clefs. The music consists of quarter and eighth notes, some with slurs and ties. The bass clef staves have square notes in the first four measures.

Fingering diagram for the first treble clef staff. It shows a sequence of chords and notes with numbers 1-5 below them. A yellow oval highlights the final chord (4, 3, 2).

3	3	3	3	3	3	3	3	3	3	4	3	2	6	6	6
3	3	3	3	3	3	3	3	3	3	4	3	2	6	6	6
3	3	3	3	3	3	3	3	3	3	4	3	2	6	6	6

```
schumann =  
corpus.parseWork('schumann/opus41no1', 3)  
schumann.plot('pitch')
```



Schumann  
Chopin

# Feature Extraction with `music21`

Why a new Feature Extraction System after jSymbolic?

Why go beyond MIDI?



## Some Elements MIDI Lacks

It does not record differences between enharmonic tones (such as E-flat and D-sharp),

It cannot precisely encode microtones (pitch bend objects being notoriously unreliable and not portable),

Few implementations support lyrics,

Differences between an eighth note and a staccato quarter is erased

Chord symbols are not supported, and so on.

[0.0, 0.0, 1.0, 0.0, 1.0, 0.0, 0.66..., 0.0, 0.0, 1.0, 0.0, 0.0]

Pitch class frequency (C, C#, D, D#/Eb, E, ...) where 1.0 is the most common pitchclass, 0.5 = half as prevalent as the most common, pitch class, etc. C-sharp is highlighted

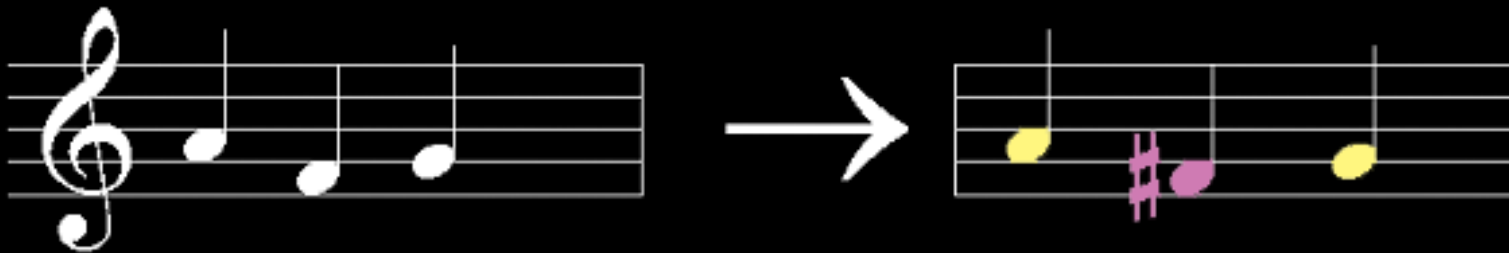
A musical score in 3/4 time with a key signature of two sharps (D major). The score consists of three staves: Treble, Middle, and Bass. The Treble staff contains a melody of quarter notes: D5, E5, F#5, G5, A5, B5, C6, D6. The Middle staff contains a chordal accompaniment of quarter notes: D4, E4, F#4, G4, A4, B4, C5, D5. The Bass staff contains a bass line of quarter notes: D3, E3, F#3, G3, A3, B3, C4, D4. The first two measures of the Treble staff and the first two measures of the Middle staff are highlighted with a yellow box. The last two measures of the Treble staff and the last two measures of the Middle staff are also highlighted with a yellow box. Below the staves, the numbers 6, 6, 6, 7 are written under the first four measures.

[0.0, 0.5, 1.0, 0.0, 0.6, 0.0, 0.4, 0.2, 0.0, 0.7, 0.0, 0.1]

After pitch realization C-sharp goes from being non-existent to one of the more prominent pitches.

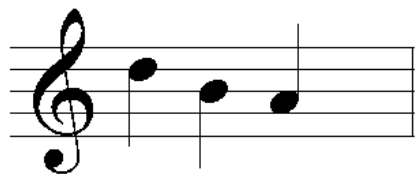
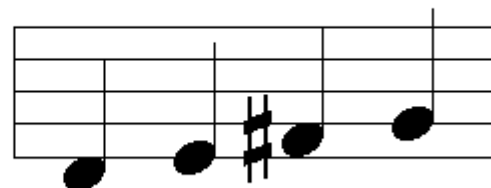
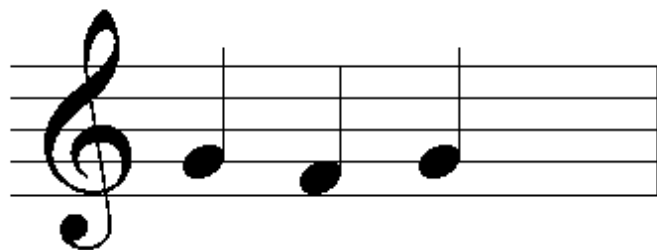
*from* The rules of “musica ficta” according to  
Nicolaus de Capua: Rule 3

*Nota quod quando cantus descendit ditonum sane et  
inmediate ascendit unam uocem. tunc de ditono  
debemus facere semiditonum.*

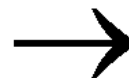
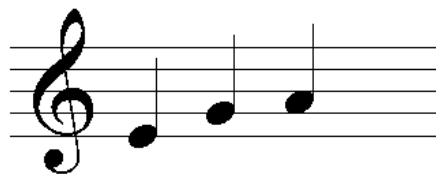


for further information, see Jan Herlinger, “Nicolaus de Capua, Antonio Zacara da Teramo, and *musica ficta*” in Francesco Zimei, editor, *Antonio Zacara da Teramo e il suo tempo* (Lucca: LIM, 2004), pp. 67–89.

# Nicolaus de Capua, the other three rules



or



58

The image shows a musical score for three voices (Soprano, Alto, and Bass) in a single system. The music is in 8/8 time, indicated by the '8' below the first staff. The lyrics are: "mus te. Glo-ri - fi-ca - mus te." for the Soprano; "te. Glo-ri fi - ca - mus te" for the Alto; and "mus te. Glo - ri - fi - ca - mus te" for the Bass. Annotations include a yellow oval around the Soprano's first two notes, a yellow oval around the Alto's notes for "fi - ca", and a yellow oval around the Bass's notes for "fi - ca - mus". A grey rectangular box highlights the Alto's notes for "fi - ca" and the Bass's notes for "fi - ca - mus".

mus te. Glo-ri - fi-ca - mus te.

te. Glo-ri fi - ca - mus te

mus te. Glo - ri - fi - ca - mus te

Courtesy of Libreria Musicale Italiana. Used with permission

Zachara, Gloria "Rosetta", PMFC 13. Herlinger, 2004,  
p. 72

```

def capuaRuleThree (stream):
    '''Applies Capua's third rule to the given stream, i.e. if there is a
    descending major third followed by an ascending major second, the second
    note will be made a half-step higher so that there is a descending minor
    third followed by an ascending minor second. Also changes
    note.editorial.color for rule 3 (green pink green).'''
    for i in range(0, len(stream.notes)-2):
        n1 = stream.notes[i]
        n2 = stream.notes[i+1]
        n3 = stream.notes[i+2]

        if n1.isRest or \
            n2.isRest or \
            n3.isRest:
            continue

        i1 = generateInterval(n1,n2)
        i2 = generateInterval(n2,n3)

        if n1.accidental is not None or \
            n2.accidental is not None or \
            n3.accidental is not None:
            continue

        ### never seems to improve things...
        if n2.step == "A" or n2.step == "D":
            continue

        # e.g., E C D => E C# D
        if i1.directedName == "M-3" and \
            i2.directedName == "M2":
            if (n2.editorial.misc.has_key("capua")):
                n2.editorial.misc['capua'] += RULETHREE
            else:
                n2.editorial.misc['capua'] = RULETHREE
            n2.editorial.ficta = Accidental("sharp")
            n2.editorial.misc["capua-ficta"] = Accidental("sharp")
            n1.editorial.color = "pink"

```

# Encoded list of incipits and cadences of Trecento ballate

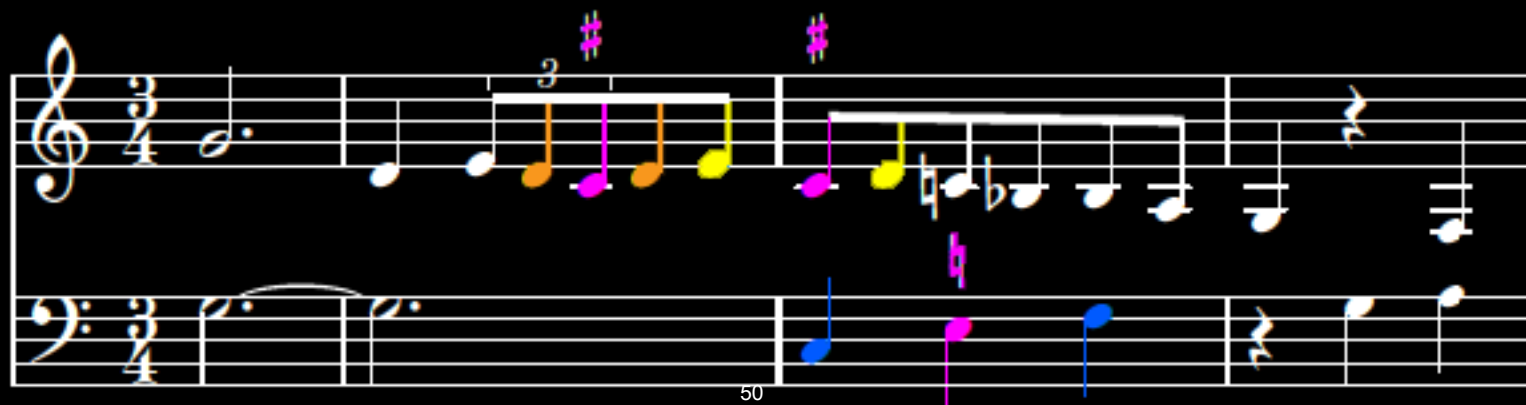
FischerNum	Incipit	pg start	pg end	begin time	incipit C	incipit T
2	Acurr' uomo	1	1	2/4	r4 a8 b c'2 ~ c' b4 r4	r4 d8 d c2 ~c4. d
3	Ad(d)io, amore mio					
12	Altro che sospirar	2	4	2/4	a2 r4 e r f# g r8 e g f e c	A2~ A A G F2~ F
16	Amor amaro					
22	Amore a lo to aspetto	5	6	2/4	A2 A8 B- c A A4 r8 A16	D2 D D
28	Amor, merçè	7	7	6/8	a4.~ a4 e8 e f d d r g	A2. c4. d4.
29	Amor mi fa cantar	8	8	9/8		
30	Amor mi stringe	9	10	2/4	g2 a4 g f16 e d8 r f	G4 B- A4. c8~ c
32	Amor, per ti sempre ardo	157	157	6/8		
39	Astio non morì may! Nè puo morir	11	12	3/4	d4 f#2 g8 f f e e d trip{d4	G4 D2 C2. F4 E-
40	A tanti homini	13	13	2/4	d2~ d4 r8 d c d e f e4 r8	G2 D8 G4 F8 E4
41	Aymè, per tutto	14	15	6/8	c4. d4.~ d4 c8 B- A16 E	F2. BB-4 F8 D C.
42	Bel fiore dança					
43	Bench'amar	158	159	3/4	g4 f8 e4 d8 e4 c r e8 f#	c2. G4 A2 A8 A C
50	Bench'l' serva	16	17	3/4	e4 f#2 g4 r8 a f g a4 a8	A2. G4 F G D F C
56	Checc' a tte piaccia	18	20	9/8	d4. e4 d8 e c d e4. f#2. b4. c4. B2. A2.~	
61	Che pensi di me	21	23	6/8	r4. r8 r c d4. r8 r d e4 g8	r4. r8 r c d4. G4.
62	Che ti çova	24	24	3/4		
72	Chi vole amar	27	28	3/4	a2.~ a2. e4 e g8 f g a g	d2.~ d2. c2. a2. g
78	Come 'nfra l'atre donne	32	33	2/4	c2 e4 f a8 g g d d4 e	c2 A4 a G2 G F B
81	Con dollia me ne vo					
83	Con lagreme sospiro	34	35	2/4	d2 d8 e16 d c8 r r d B- c	G2 G4 A8 B- G E
88	Cosa non è ch'a se tanto	29	31	9/8	c2. d4. e4. d4 e8 c4 d8	C2.~ C4. G2. G4
97	Deduto sey a quel	36	40	3/4	g2. a2 g4 a trip{g4 a8} f	c2. c4 c2 F2. c4



# Capua rules applied to *Non creder, donna* by Francesco da Firenze (Landini)

The image displays two systems of musical notation, each consisting of two staves. The notation is in bass clef with a 2/4 time signature. The first system shows the original notation. The second system highlights specific notes with colored boxes and sharp signs (#) above them, illustrating the application of the Capua rules. In the first system of the second system, the notes G4 and A4 in the upper staff are highlighted in yellow, and the note G4 in the lower staff is highlighted in pink. In the second system of the second system, the notes G4 and A4 in the upper staff are highlighted in yellow, and the notes G4 and A4 in the lower staff are highlighted in green. Sharp signs (#) are placed above the highlighted notes in both systems of the second system.

But does it *always* work?



# Does “capua” fix aug./dim. unisons, 5ths, & 8ves?

Augmented/Diminished 1/5/8 that “capua” left alone:

269

Augmented/Diminished 1/5/8 that “capua” fixed:

26 (9%)

Perfect consonances that “capua” left alone: 11522

Perfect consonances that “capua” made worse: 287 (2.4%)

(Is it better than nothing 3 out of 4 times?

or worse than nothing 9 out of 10 times?)

# Reuse your repertory

Exploring tonal closure in ballate

# Francesco da Firenze (or 'Landini') *image from the Squarcialupi Codex*



# Francesco's Tonality: Summary

## *Francesco Cadences*

<b>Total Same</b>	<b>64</b>	<b>46.4%</b>
<b>Total Different</b>	<b>74</b>	<b>53.6%</b>

## *Other Attributed Composers Cadences*

<b>Total Same</b>	<b>49</b>	<b>32.0%</b>
<b>Total Different</b>	<b>104</b>	<b>68.0%</b>

## *Anonymous Cadences*

<b>Total Same</b>	<b>28</b>	<b>37.8%</b>
<b>Total Different</b>	<b>46</b>	<b>62.2%</b>

# Francesco's Tonal usage

Tenor notes at the beginning and first cadences of ballate

<b>1st note</b>	<b>Cadence note</b>	<b>#</b>
<b>A</b>	<b>same</b>	<b>0</b>
	<b>different</b>	<b>16</b>
	<i>total</i>	<i>16 (12% of total)</i>
<b>C</b>	<b>same</b>	<b>12</b>
	<b>different</b>	<b>7</b>
	<i>total</i>	<i>19 (14%)</i>
<b>D</b>	<b>same</b>	<b>25</b>
	<b>different</b>	<b>7</b>
	<i>total</i>	<i>32 (23%)</i>

<b>1st note</b>	<b>Cadence note</b>	<b>#</b>	
<b>E</b>	<b>same</b>	<b>0</b>	
	<b>diff.</b>	<b>2</b>	
	<i>total</i>	<i>2</i>	<i>( 1%)</i>
<b>F</b>	<b>same</b>	<b>8</b>	
	<b>diff.</b>	<b>8</b>	
	<i>total</i>	<i>16</i>	<i>(12%)</i>
<b>G</b>	<b>same</b>	<b>19</b>	
	<b>diff.</b>	<b>34</b>	
	<i>total</i>	<i>53</i>	<i>(38%)</i>



Making a corpus from a single example

A Canon by Johannes Ciconia

J. Cicoma.

Handwritten musical score on four staves. The notation consists of diamond-shaped notes with stems, typical of early printed music. The lyrics are written in black ink below the staves. A red annotation is written across the second and third staves. The music is in a key with one sharp (F#) and a common time signature (C). The lyrics are: "quod factatur a iuribus opere non demonstratur. Ut aqua pennis sepius scientia de ne ga tur." The red annotation reads: "Tenor quem contra tenor. triplūq; fugant temporibus in quinq;".

quod factatur a iuribus opere non demonstratur. Ut aqua pennis  
*Tenor quem contra tenor. triplūq; fugant temporibus in quinq;*  
sepius scientia de ne ga tur.

```

185     return qjPart
186
187 def findRetrogradeVoices(show = True):
188     '''
189     the structure of the piece strongly suggests a retrograde solution
190     (e.g., there is a cadence in m5 and five measures from the end and one
191     at the exact center). This method tries all transpositions of one
192     voice vs. the other and gives positive points to intervals of 3, 4,
193     5, 6, and 8 (incl. tritones, since they might be fixed w/ other voices;
194     4th is included since there could be a 3rd or 5th below it).
195     '''
196
197     for transpose in [1, 2, -2, 3, -3, 4, -4]:
198         for invert in [False, True]:
199             qj1 = getQJ()
200             qj2 = getQJ()
201             if transpose != 1:
202                 transposeStreamDiatonic(qj2, transpose)
203             if invert is True:
204                 qj2.invertDiatonic(qj2.flat.notes[0], inplace = True)
205             qj2 = reverse(qj2, makeNotation = False)
206             qj = stream.Score()
207             qj.insert(0, qj2.flat)
208             qj.insert(0, qj1.flat)
209             qjChords = qj.chordify()
210             consScore = 0
211             totIntervals = 0
212             for n in qjChords.flat.notes:

```

Trans: 2  
Invert: True

Score: 134

2.0 0.25 -2.0 -0.5 2.0 0.25 1.0 -0.5 4.0 2.0 4.0 2.0 -4.0 0.25 -2.0 0.25 32 -4.0 0.25 -2.0 0.25

2.0 0.25 1.0 0.25 2.0 -0.5 1.0 0.25 2.0 0.25 2.0 2.0 0.25 1.0 0.25 0.25 1.0 0.25 2.0 -0.5 1.0 0.25 2.0 0.25 1.0 0.25 -4.0 0.25 2.0

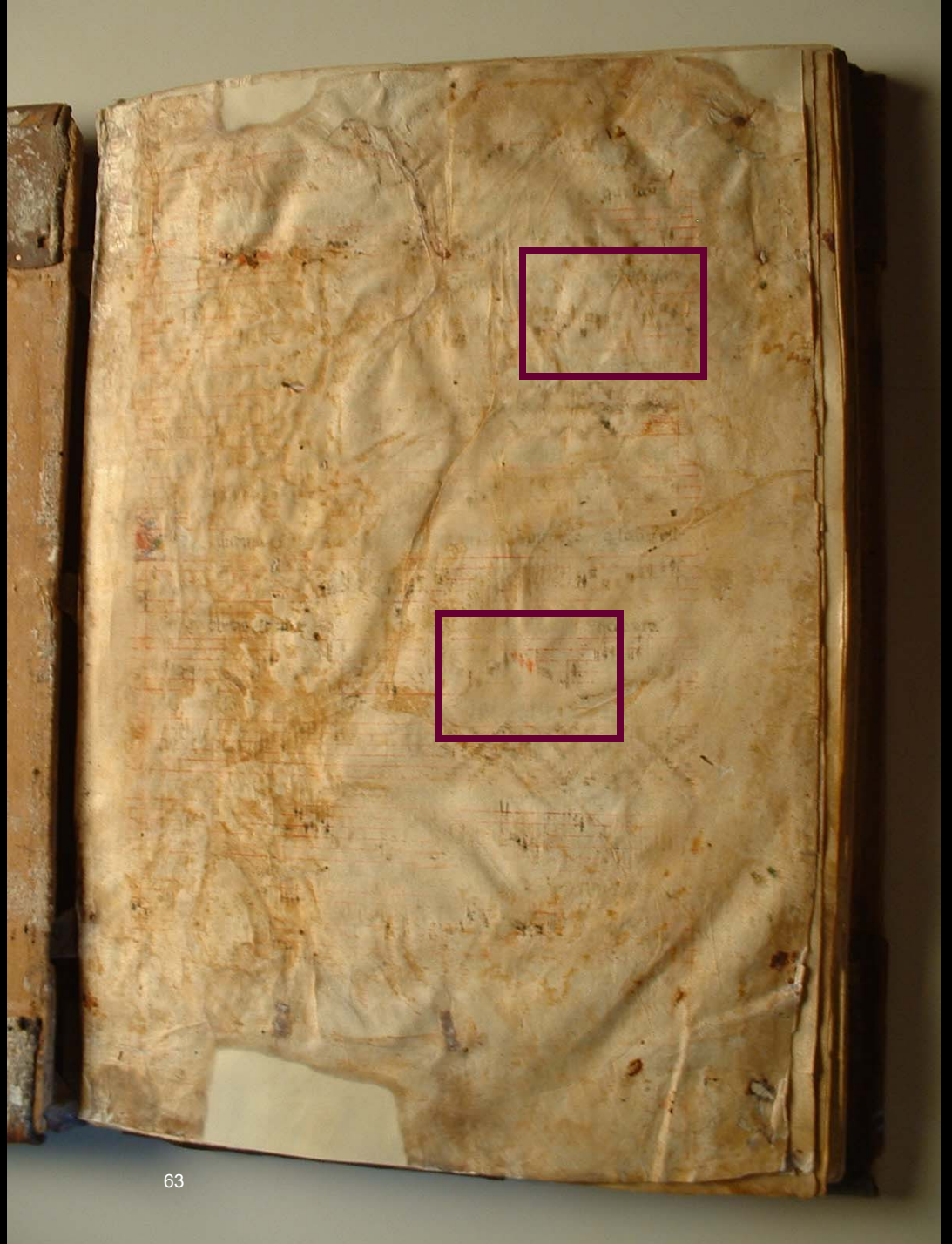
# Simultaneous retrograde solutions for Quod Jactatur

INTERVAL	INVERT?	SCORE
1		113
1	Invert	-150
2		-111
2	Invert	134
-2		-119
-2	Invert	90
3		121
3	Invert	-110
-3		103
-3	Invert	92
4		112
4	Invert	123
-4		123
-4	Invert	125

# Short case studies

## 3. Searching in Medieval and Renaissance Music

Cividale 98, folio 1r

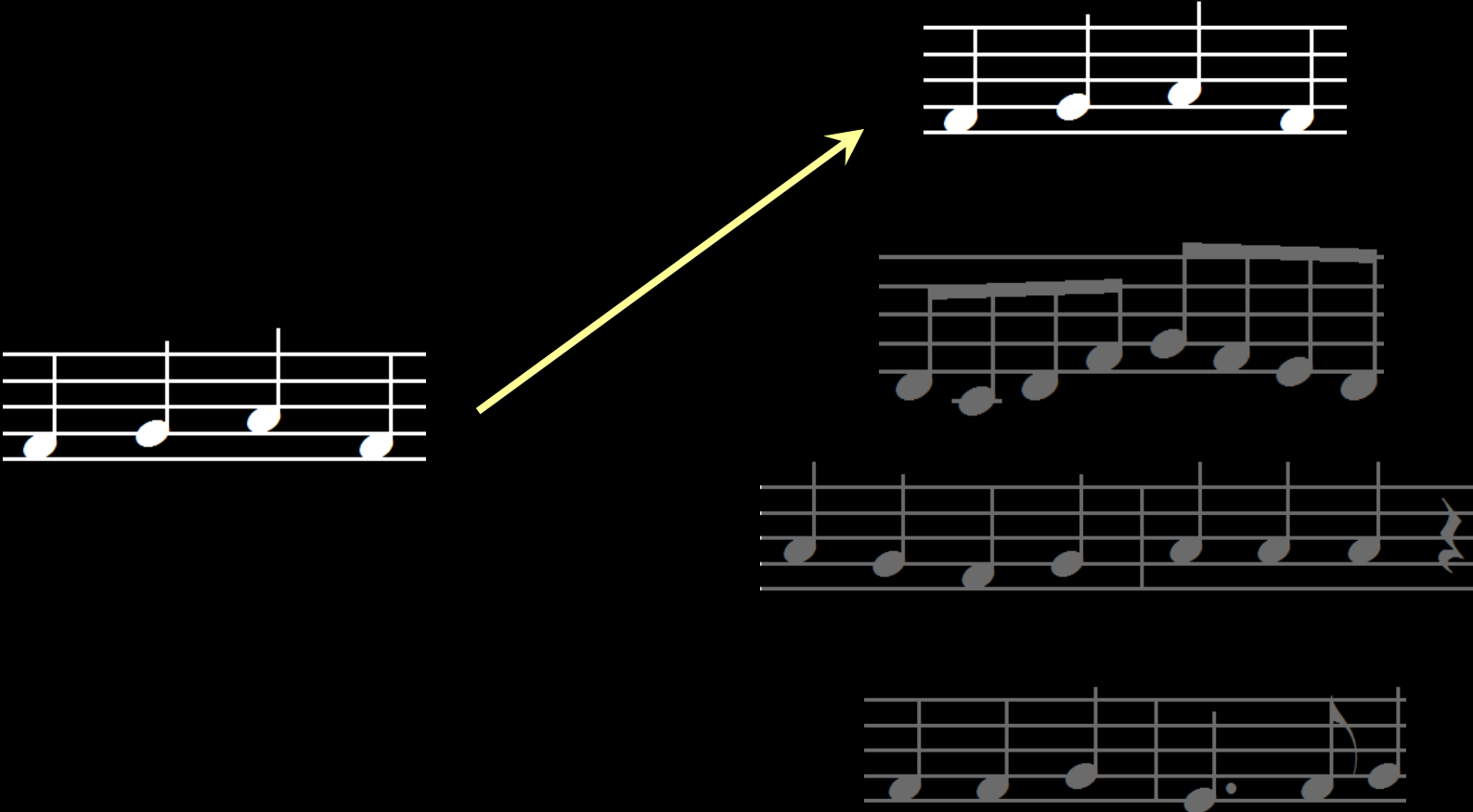


# Cividale 98, digitally enhanced

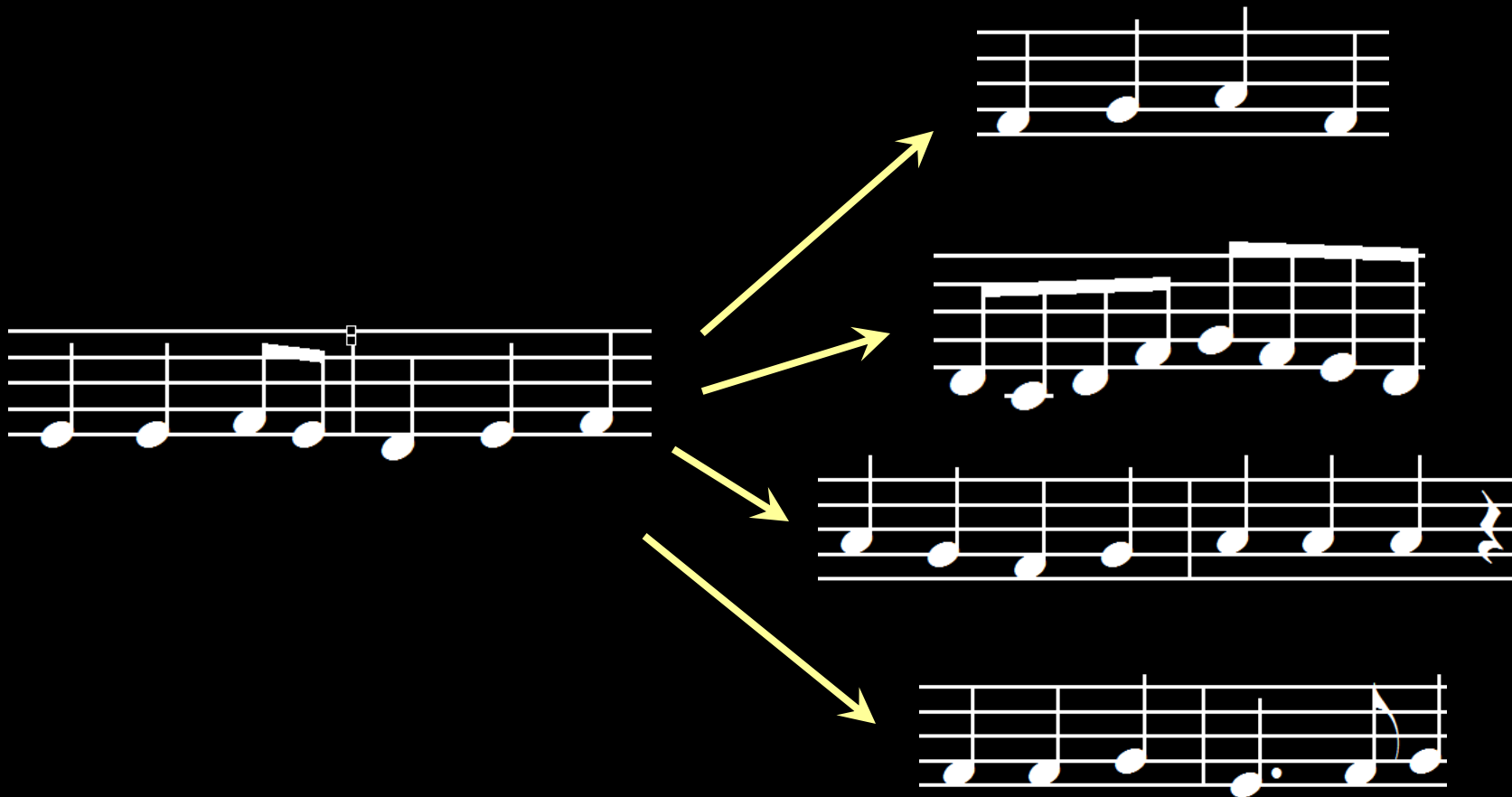




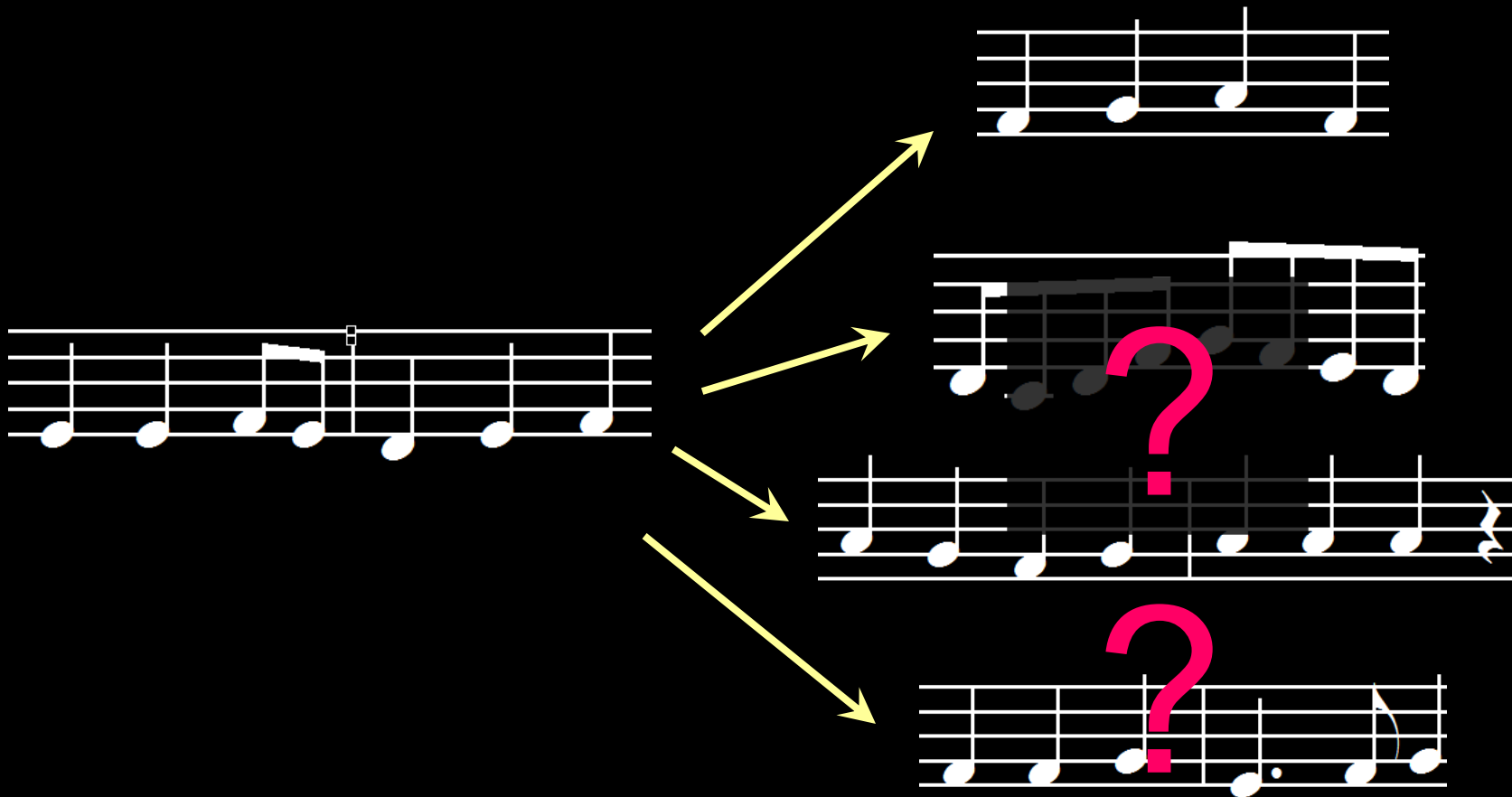
# Exact matches are no problem for current databases



But even the smallest musical changes pose problems



But even the smallest musical changes pose problems



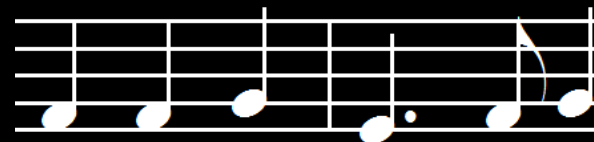
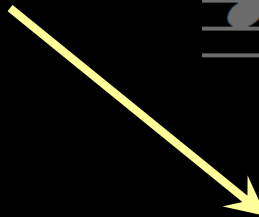
# music21 encodes information that solves these search problems.

sometimes melodies are higher or lower than in the database



rhythms change slightly

“skips” get filled in



## Code: Music21 for searching on Cividale 98 piece

```
score.genre == "Ballade" or score.genre in  
    medRen.MassMovements
```

```
score.manuscript.notation["red"] is not False
```

```
score.timeSignature in ["6/8", "9/8"]
```

```
for thisPart in score:
```

```
    thisPart.notes.containsIntervals(-2, -2, -2)
```

```
    search.notes.thisPart("r8 r8 ?8 ?4. * r8 r8 ?4.") is not None
```



*This search has been simplified slightly for rapid presentation.*

# Cividale 98, digitally enhanced



Cividale 98, digitally enhanced and identified as a Credo by Zachara da Teramo



T: mm. 321-329 (+custos of 330)



Ct: mm. 316-323



## Creating a Reduction and Labeling Intervals

This example, after parsing a polyphonic work stored as an Opus and creating a Score, presents and labels the intervals of the resultant chords of all distinct harmonies:

```
from music21 import corpus
# Parse an Opus, a collection of Scores
o = corpus.parsework('josquin/laDeplorationDeLaMorteDeJohannesOckeghem')
# Create a Score from a Measure range
sExcerpt = o.mergeScores().measures(127, 134)
# Create a reduction of Chords
reduction = sExcerpt.chordify()
# Iterate over the Chords and prepare presentation
for c in reduction.flat.getElementsByClass('Chord'):
    c.closedPosition(forceOctave=4, inPlace=True)
    c.removeRedundantPitches(inPlace=True)
    c.annotateIntervals()
# Add the reduction and display the results
sExcerpt.insert(0, reduction)
sExcerpt.show()
```



- Quantity
- Searching the Corpus by Locale
- Finding Chords by Root and Collecting their Successors
- [Pitch and Duration Transformations](#)
- Basic Counting of and Searching for Musical Elements

Previous topic

Overview: Meters, Time Signatures, and Processing Beams, Accents, and Beats

Next topic

Installing Music21

This Page

Show Source

Quick search





## Short case studies

### 4. Producing digital and print editions in different styles



E in terra pax hominibus bone voluntatis laudam te benedicim  
 te Adoramus te glorificamus te gratias agimus tibi propter magnam gloriam tuam Dne  
 deus pater celestis deus pater omnipotens dne fili vnguentis ihu xpc dne deus agnus dei  
 filius patris Qui tollis peccata mundi miserere nobis qui tollis peccata mundi  
 suscipe deprecationes nostras Qui sedes ad dexteram patris miserere nobis quoniam  
 tu solus sanctus tu solus dnus tu solus altissimus ihu xpc  
 gloria dei pa tris  
 amen **E** in terra pax laudam te



Digital Image Archive of Medieval Music

from music21 import \*

gloria = corpus.parse('luca/gloria')

gloria.show()

The image displays a musical score for a Gloria, specifically the section 'Et in terra pax hominibus bonae voluntatis'. The score is presented in a standard musical notation format with three systems of staves. Each system consists of a vocal line (Soprano, Contratenor, or Tenor) and a piano accompaniment line. The lyrics are written below the vocal lines. The first system shows the beginning of the piece with the lyrics 'Et in terra pax hominibus bonae'. The second system continues with 'voluntatis. Laudamus te. Benedicimus te.' The third system shows the continuation of the piano accompaniment. The score is written in 6/8 time and features a variety of rhythmic values and melodic lines.

Et in terra pax ho - mi - ni - bus bo - ne

Contratenor et in terra pax hominibus

Tenor Et in terra pax

vo - lun - ta - tis. Lau - da - mus te. Be - ne - di - ci - mus te.

Laudamus te

```
from music21 import *
```

```
gloria = corpus.parse('luca/gloria')
```

```
gloriaNew = medren.convertHouseStyle(gloria, durationScale = 2,  
barlineStyle = 'tick', tieTransfer = True)
```

```
gloriaNew.show()
```

Et in terra pax ho - mi - ni - bus bo - ne

Contratenor et in terra pax hominibus

Tenor Et in terra pax

8  
vo - lun - ta - tis. Lau - da - mus

Laudamus te

Laudamus te

# Counting our Losses

Two questions about Medieval music:

- (1) How many **manuscripts** of medieval polyphony were once copied but no longer exist?
- (2) How many **pieces** of medieval polyphony were once copied but no longer exist?

## Counting our Losses

What percentage of **pieces** of 14th c. Italian polyphony were once copied but no longer exist?

- (a) 90% of them?
- (b) 99% of them?
- (c) 99.9% of them?
- (d) 99.9999% of them?

How could we figure out the answer?

$$N \times P(m) = m$$

number of pieces originally      probability that any given piece is missing      total # missing pieces



How can we find  $P(m)$ ?  
[ probability that we're missing a given piece ]

What's the probability that any given piece is in a given manuscript?

$$\frac{r_i}{N}$$

$r_i$  = number of pieces in that manuscript

$N$  = total number of pieces originally copied.

[a number we don't know...]

\* unquestioned assumption: pieces have equal probability of being copied.

How can we find  $P(m)$ ?  
[ probability that we're missing a given piece ]

What's the probability that any given piece is **NOT** in a given manuscript?

$$1 - \frac{r_i}{N}$$

\* unquestioned assumption still in play.

How can we find  $P(m)$ ?  
[ probability that we're missing a given piece ]

What's the probability that any given piece is **NOT** in  
**TWO** given manuscripts?

$$\left(1 - \frac{r_1}{N}\right) \left(1 - \frac{r_2}{N}\right)$$

\* unquestioned assumption still in play.

# How can we find P(m)?

[ probability that we're missing a given piece ]

What's the probability that any given piece is **NOT** in **ANY** of all 85 surviving manuscripts?

$$\begin{aligned} & \left(1 - \frac{r_1}{N}\right) \left(1 - \frac{r_2}{N}\right) \left(1 - \frac{r_3}{N}\right) \left(1 - \frac{r_4}{N}\right) \left(1 - \frac{r_5}{N}\right) \left(1 - \frac{r_6}{N}\right) \left(1 - \frac{r_7}{N}\right) \left(1 - \frac{r_8}{N}\right) \left(1 - \frac{r_9}{N}\right) \left(1 - \frac{r_{10}}{N}\right) \left(1 - \frac{r_{11}}{N}\right) \left(1 - \frac{r_{12}}{N}\right) \left(1 - \frac{r_{13}}{N}\right) \dots \\ & \left(1 - \frac{r_{14}}{N}\right) \left(1 - \frac{r_{15}}{N}\right) \left(1 - \frac{r_{16}}{N}\right) \left(1 - \frac{r_{17}}{N}\right) \left(1 - \frac{r_{18}}{N}\right) \left(1 - \frac{r_{19}}{N}\right) \left(1 - \frac{r_{20}}{N}\right) \left(1 - \frac{r_{21}}{N}\right) \left(1 - \frac{r_{22}}{N}\right) \left(1 - \frac{r_{23}}{N}\right) \left(1 - \frac{r_{24}}{N}\right) \dots \\ & \left(1 - \frac{r_{25}}{N}\right) \left(1 - \frac{r_{26}}{N}\right) \left(1 - \frac{r_{27}}{N}\right) \left(1 - \frac{r_{28}}{N}\right) \left(1 - \frac{r_{29}}{N}\right) \left(1 - \frac{r_{30}}{N}\right) \left(1 - \frac{r_{31}}{N}\right) \left(1 - \frac{r_{32}}{N}\right) \left(1 - \frac{r_{33}}{N}\right) \left(1 - \frac{r_{34}}{N}\right) \left(1 - \frac{r_{35}}{N}\right) \dots \\ & \left(1 - \frac{r_{36}}{N}\right) \left(1 - \frac{r_{37}}{N}\right) \left(1 - \frac{r_{38}}{N}\right) \left(1 - \frac{r_{39}}{N}\right) \left(1 - \frac{r_{40}}{N}\right) \left(1 - \frac{r_{41}}{N}\right) \left(1 - \frac{r_{42}}{N}\right) \left(1 - \frac{r_{43}}{N}\right) \left(1 - \frac{r_{44}}{N}\right) \left(1 - \frac{r_{45}}{N}\right) \left(1 - \frac{r_{46}}{N}\right) \dots \\ & \left(1 - \frac{r_{47}}{N}\right) \left(1 - \frac{r_{48}}{N}\right) \left(1 - \frac{r_{49}}{N}\right) \left(1 - \frac{r_{50}}{N}\right) \left(1 - \frac{r_{51}}{N}\right) \left(1 - \frac{r_{52}}{N}\right) \left(1 - \frac{r_{53}}{N}\right) \left(1 - \frac{r_{54}}{N}\right) \left(1 - \frac{r_{55}}{N}\right) \left(1 - \frac{r_{56}}{N}\right) \left(1 - \frac{r_{57}}{N}\right) \dots \\ & \left(1 - \frac{r_{58}}{N}\right) \left(1 - \frac{r_{59}}{N}\right) \left(1 - \frac{r_{60}}{N}\right) \left(1 - \frac{r_{61}}{N}\right) \left(1 - \frac{r_{62}}{N}\right) \left(1 - \frac{r_{63}}{N}\right) \left(1 - \frac{r_{64}}{N}\right) \left(1 - \frac{r_{65}}{N}\right) \left(1 - \frac{r_{66}}{N}\right) \left(1 - \frac{r_{67}}{N}\right) \left(1 - \frac{r_{68}}{N}\right) \dots \\ & \left(1 - \frac{r_{69}}{N}\right) \left(1 - \frac{r_{70}}{N}\right) \left(1 - \frac{r_{71}}{N}\right) \left(1 - \frac{r_{72}}{N}\right) \left(1 - \frac{r_{73}}{N}\right) \left(1 - \frac{r_{74}}{N}\right) \left(1 - \frac{r_{75}}{N}\right) \left(1 - \frac{r_{76}}{N}\right) \left(1 - \frac{r_{77}}{N}\right) \left(1 - \frac{r_{78}}{N}\right) \left(1 - \frac{r_{79}}{N}\right) \dots \\ & \left(1 - \frac{r_{80}}{N}\right) \left(1 - \frac{r_{81}}{N}\right) \left(1 - \frac{r_{82}}{N}\right) \left(1 - \frac{r_{83}}{N}\right) \left(1 - \frac{r_{84}}{N}\right) \left(1 - \frac{r_{85}}{N}\right) = P(m) \end{aligned}$$

\* unquestioned assumption still in play.

How could we figure out the answer?

$$N \times P(m) = m$$

number of  
pieces originally

probability that  
any given piece  
is missing

total # missing  
pieces

\* unquestioned assumption still in play.

# How could we figure out the answer?

$$\begin{array}{l}
 \mathbf{N} \times \\
 \text{number of} \\
 \text{pieces originally}
 \end{array}
 \left(1 - \frac{r_1}{N}\right) \left(1 - \frac{r_2}{N}\right) \left(1 - \frac{r_3}{N}\right) \left(1 - \frac{r_4}{N}\right) \left(1 - \frac{r_5}{N}\right) \left(1 - \frac{r_6}{N}\right) \left(1 - \frac{r_7}{N}\right) \left(1 - \frac{r_8}{N}\right) \left(1 - \frac{r_9}{N}\right) \left(1 - \frac{r_{10}}{N}\right) \left(1 - \frac{r_{11}}{N}\right) \left(1 - \frac{r_{12}}{N}\right) \left(1 - \frac{r_{13}}{N}\right) \dots$$

$$\begin{array}{l}
 \left(1 - \frac{r_{14}}{N}\right) \left(1 - \frac{r_{15}}{N}\right) \left(1 - \frac{r_{16}}{N}\right) \left(1 - \frac{r_{17}}{N}\right) \left(1 - \frac{r_{18}}{N}\right) \left(1 - \frac{r_{19}}{N}\right) \left(1 - \frac{r_{20}}{N}\right) \left(1 - \frac{r_{21}}{N}\right) \left(1 - \frac{r_{22}}{N}\right) \left(1 - \frac{r_{23}}{N}\right) \left(1 - \frac{r_{24}}{N}\right) \dots \\
 \left(1 - \frac{r_{25}}{N}\right) \left(1 - \frac{r_{26}}{N}\right) \left(1 - \frac{r_{27}}{N}\right) \left(1 - \frac{r_{28}}{N}\right) \left(1 - \frac{r_{29}}{N}\right) \left(1 - \frac{r_{30}}{N}\right) \left(1 - \frac{r_{31}}{N}\right) \left(1 - \frac{r_{32}}{N}\right) \left(1 - \frac{r_{33}}{N}\right) \left(1 - \frac{r_{34}}{N}\right) \left(1 - \frac{r_{35}}{N}\right) \dots \\
 \left(1 - \frac{r_{36}}{N}\right) \left(1 - \frac{r_{37}}{N}\right) \left(1 - \frac{r_{38}}{N}\right) \left(1 - \frac{r_{39}}{N}\right) \left(1 - \frac{r_{40}}{N}\right) \left(1 - \frac{r_{41}}{N}\right) \left(1 - \frac{r_{42}}{N}\right) \left(1 - \frac{r_{43}}{N}\right) \left(1 - \frac{r_{44}}{N}\right) \left(1 - \frac{r_{45}}{N}\right) \left(1 - \frac{r_{46}}{N}\right) \dots \\
 \left(1 - \frac{r_{47}}{N}\right) \left(1 - \frac{r_{48}}{N}\right) \left(1 - \frac{r_{49}}{N}\right) \left(1 - \frac{r_{50}}{N}\right) \left(1 - \frac{r_{51}}{N}\right) \left(1 - \frac{r_{52}}{N}\right) \left(1 - \frac{r_{53}}{N}\right) \left(1 - \frac{r_{54}}{N}\right) \left(1 - \frac{r_{55}}{N}\right) \left(1 - \frac{r_{56}}{N}\right) \left(1 - \frac{r_{57}}{N}\right) \dots \\
 \left(1 - \frac{r_{58}}{N}\right) \left(1 - \frac{r_{59}}{N}\right) \left(1 - \frac{r_{60}}{N}\right) \left(1 - \frac{r_{61}}{N}\right) \left(1 - \frac{r_{62}}{N}\right) \left(1 - \frac{r_{63}}{N}\right) \left(1 - \frac{r_{64}}{N}\right) \left(1 - \frac{r_{65}}{N}\right) \left(1 - \frac{r_{66}}{N}\right) \left(1 - \frac{r_{67}}{N}\right) \left(1 - \frac{r_{68}}{N}\right) \dots \\
 \left(1 - \frac{r_{69}}{N}\right) \left(1 - \frac{r_{70}}{N}\right) \left(1 - \frac{r_{71}}{N}\right) \left(1 - \frac{r_{72}}{N}\right) \left(1 - \frac{r_{73}}{N}\right) \left(1 - \frac{r_{74}}{N}\right) \left(1 - \frac{r_{75}}{N}\right) \left(1 - \frac{r_{76}}{N}\right) \left(1 - \frac{r_{77}}{N}\right) \left(1 - \frac{r_{78}}{N}\right) \left(1 - \frac{r_{79}}{N}\right) \dots \\
 \left(1 - \frac{r_{80}}{N}\right) \left(1 - \frac{r_{81}}{N}\right) \left(1 - \frac{r_{82}}{N}\right) \left(1 - \frac{r_{83}}{N}\right) \left(1 - \frac{r_{84}}{N}\right) \left(1 - \frac{r_{85}}{N}\right)
 \end{array}
 = m$$

total # missing pieces

\* unquestioned assumption still in play.

What's the number of missing pieces?

$$m = N - S$$

# of missing pieces                      # of pieces originally                      # of pieces that survive

\* unquestioned assumption still in play.

# How could we figure out the answer?

$$\begin{array}{l}
 \mathbf{N} \times \\
 \# \text{ original} \\
 \text{pieces}
 \end{array}
 \left(1 - \frac{r_1}{N}\right) \left(1 - \frac{r_2}{N}\right) \left(1 - \frac{r_3}{N}\right) \left(1 - \frac{r_4}{N}\right) \left(1 - \frac{r_5}{N}\right) \left(1 - \frac{r_6}{N}\right) \left(1 - \frac{r_7}{N}\right) \left(1 - \frac{r_8}{N}\right) \left(1 - \frac{r_9}{N}\right) \left(1 - \frac{r_{10}}{N}\right) \left(1 - \frac{r_{11}}{N}\right) \left(1 - \frac{r_{12}}{N}\right) \left(1 - \frac{r_{13}}{N}\right) \dots$$

$$\begin{array}{l}
 \left(1 - \frac{r_{14}}{N}\right) \left(1 - \frac{r_{15}}{N}\right) \left(1 - \frac{r_{16}}{N}\right) \left(1 - \frac{r_{17}}{N}\right) \left(1 - \frac{r_{18}}{N}\right) \left(1 - \frac{r_{19}}{N}\right) \left(1 - \frac{r_{20}}{N}\right) \left(1 - \frac{r_{21}}{N}\right) \left(1 - \frac{r_{22}}{N}\right) \left(1 - \frac{r_{23}}{N}\right) \left(1 - \frac{r_{24}}{N}\right) \dots \\
 \left(1 - \frac{r_{25}}{N}\right) \left(1 - \frac{r_{26}}{N}\right) \left(1 - \frac{r_{27}}{N}\right) \left(1 - \frac{r_{28}}{N}\right) \left(1 - \frac{r_{29}}{N}\right) \left(1 - \frac{r_{30}}{N}\right) \left(1 - \frac{r_{31}}{N}\right) \left(1 - \frac{r_{32}}{N}\right) \left(1 - \frac{r_{33}}{N}\right) \left(1 - \frac{r_{34}}{N}\right) \left(1 - \frac{r_{35}}{N}\right) \dots \\
 \left(1 - \frac{r_{36}}{N}\right) \left(1 - \frac{r_{37}}{N}\right) \left(1 - \frac{r_{38}}{N}\right) \left(1 - \frac{r_{39}}{N}\right) \left(1 - \frac{r_{40}}{N}\right) \left(1 - \frac{r_{41}}{N}\right) \left(1 - \frac{r_{42}}{N}\right) \left(1 - \frac{r_{43}}{N}\right) \left(1 - \frac{r_{44}}{N}\right) \left(1 - \frac{r_{45}}{N}\right) \left(1 - \frac{r_{46}}{N}\right) \dots \\
 \left(1 - \frac{r_{47}}{N}\right) \left(1 - \frac{r_{48}}{N}\right) \left(1 - \frac{r_{49}}{N}\right) \left(1 - \frac{r_{50}}{N}\right) \left(1 - \frac{r_{51}}{N}\right) \left(1 - \frac{r_{52}}{N}\right) \left(1 - \frac{r_{53}}{N}\right) \left(1 - \frac{r_{54}}{N}\right) \left(1 - \frac{r_{55}}{N}\right) \left(1 - \frac{r_{56}}{N}\right) \left(1 - \frac{r_{57}}{N}\right) \dots \\
 \left(1 - \frac{r_{58}}{N}\right) \left(1 - \frac{r_{59}}{N}\right) \left(1 - \frac{r_{60}}{N}\right) \left(1 - \frac{r_{61}}{N}\right) \left(1 - \frac{r_{62}}{N}\right) \left(1 - \frac{r_{63}}{N}\right) \left(1 - \frac{r_{64}}{N}\right) \left(1 - \frac{r_{65}}{N}\right) \left(1 - \frac{r_{66}}{N}\right) \left(1 - \frac{r_{67}}{N}\right) \left(1 - \frac{r_{68}}{N}\right) \dots \\
 \left(1 - \frac{r_{69}}{N}\right) \left(1 - \frac{r_{70}}{N}\right) \left(1 - \frac{r_{71}}{N}\right) \left(1 - \frac{r_{72}}{N}\right) \left(1 - \frac{r_{73}}{N}\right) \left(1 - \frac{r_{74}}{N}\right) \left(1 - \frac{r_{75}}{N}\right) \left(1 - \frac{r_{76}}{N}\right) \left(1 - \frac{r_{77}}{N}\right) \left(1 - \frac{r_{78}}{N}\right) \left(1 - \frac{r_{79}}{N}\right) \dots \\
 \left(1 - \frac{r_{80}}{N}\right) \left(1 - \frac{r_{81}}{N}\right) \left(1 - \frac{r_{82}}{N}\right) \left(1 - \frac{r_{83}}{N}\right) \left(1 - \frac{r_{84}}{N}\right) \left(1 - \frac{r_{85}}{N}\right)
 \end{array}
 = \mathbf{N} - \mathbf{S}$$

$$\begin{array}{l}
 \# \text{ original} \\
 \text{pieces}
 \end{array}
 \qquad
 \begin{array}{l}
 \# \text{ original} \\
 \text{pieces}
 \end{array}
 \qquad
 \begin{array}{l}
 \# \text{ surviving} \\
 \text{pieces}
 \end{array}$$

\* unquestioned assumption still in play.



```

#!/usr/local/bin/perl

##### find_n.pl -- Michael Scott Cuthbert
### Find hypothetical total number of pieces given X1 pieces randomly
### distributed in manuscripts of size N1 N2 N3 N4...

###    ./find_n.pl X1 N1 N2 N3 N4 ...

use strict;
use Math::BigFloat;

my $pieces_surviving_today = shift @ARGV;
my @ms_sizes    = @ARGV;
my $total_number_of_mss = scalar @ms_sizes;

#  $n * (1/n^{(\text{num\_of\_mss})}) * (n - a_1) * (n - a_2) * \dots * (n - a_y) = n - \text{pieces\_surviving\_today}$  (r)

    ### n = our current guess for the number of original pieces; start by
    ###    supposing we have them all (plus 1 to avoid division by zero).
my $n = $pieces_surviving_today + 1;

```

\* unquestioned assumption still in play.

Testing whether our model of “random copying” messes up the number for N.

Holdout cross-validation:

Look at a subset of the data (say all manuscripts discovered before 1955) and see how many pieces we would expect to have today given the new manuscripts discovered since then. And then we compare that number to what we actually have.

Result:

Pretty good! Estimates are about 4% too low.

# Original numbers for pieces in MSS 1380-1420

	# today	estimated original total	% missing
Ballate	409	507	19 %
Madrigals	167	177	6 %
Liturgical Pieces	116	196	41 %
Non-Liturgical Latin works	47	105	55 %
<b>Total</b>	<b>739</b>	<b>985</b>	<b>25 %</b>

## Counting our Losses

How many **pieces** of 14th c. Italian polyphony were once copied but no longer exist?

- (a) 90% of them?
- (b) 99% of them?
- (c) 99.9% of them?
- (d) 99.9999% of them?
- (e) **25% of them!**

MIT OpenCourseWare  
<http://ocw.mit.edu>

21M.269 Studies in Western Music History: Quantitative and Computational Approaches to Music History  
Spring 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.