Representing and Linking Music Performance Data with Score Information

Johanna Devaney and Hubert Léveillé Gauvin
School of Music, The Ohio State University

Introduction

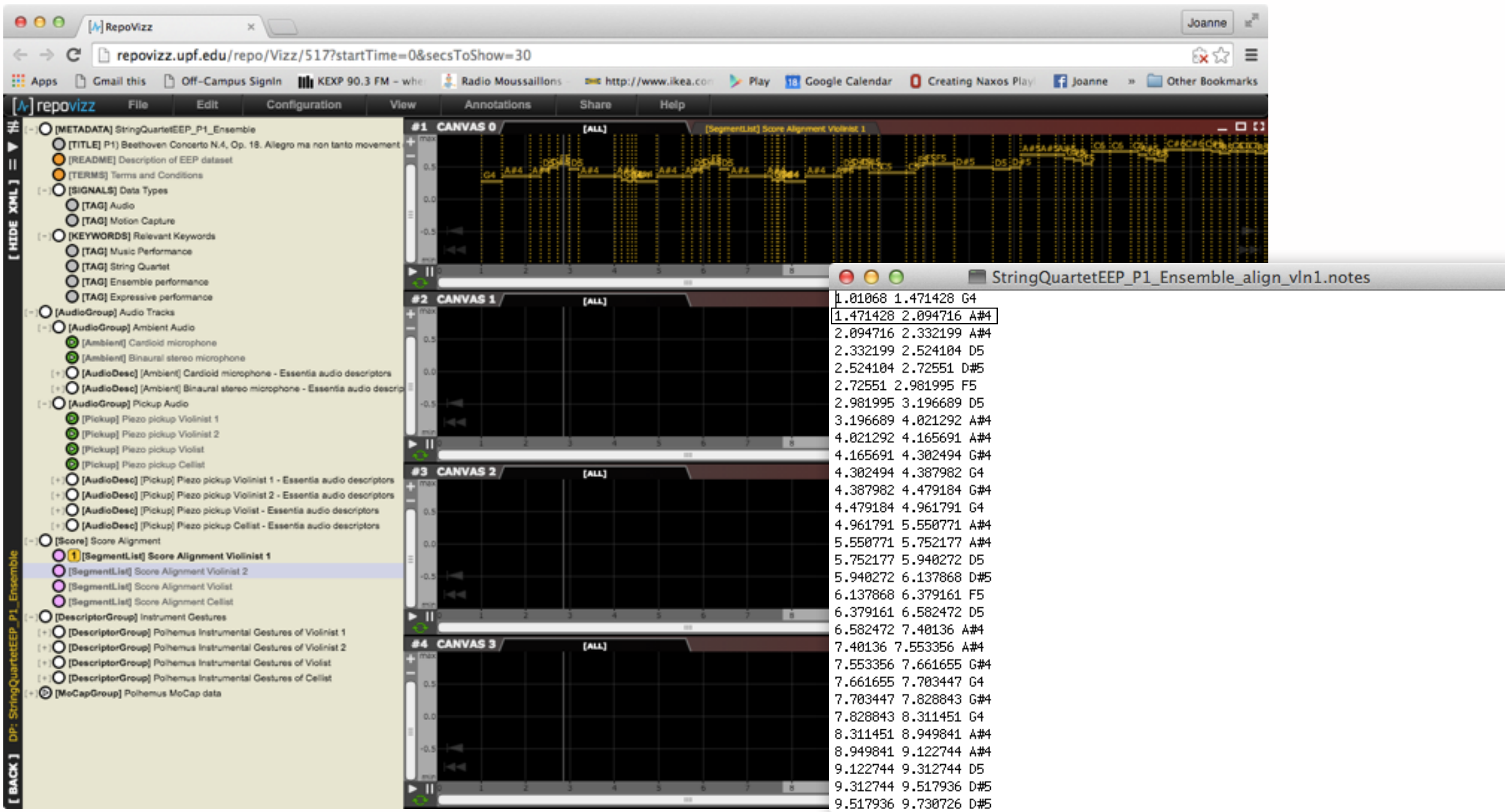
Storing and sharing performance data

- **Goal: Make performance data available to music scholars through encodings that explicitly connect score information to performance data**
- **Motivation: Increased interest by musicologists in studying performance (e.g. Cook, 2014) and recent developments for automatically estimating performance parameters (e.g., Cannam et al., 2006; Devaney, Mandel and Fujinaga, 2012; Gingras and McAdams, 2011)**

Current encoding formats

Note level

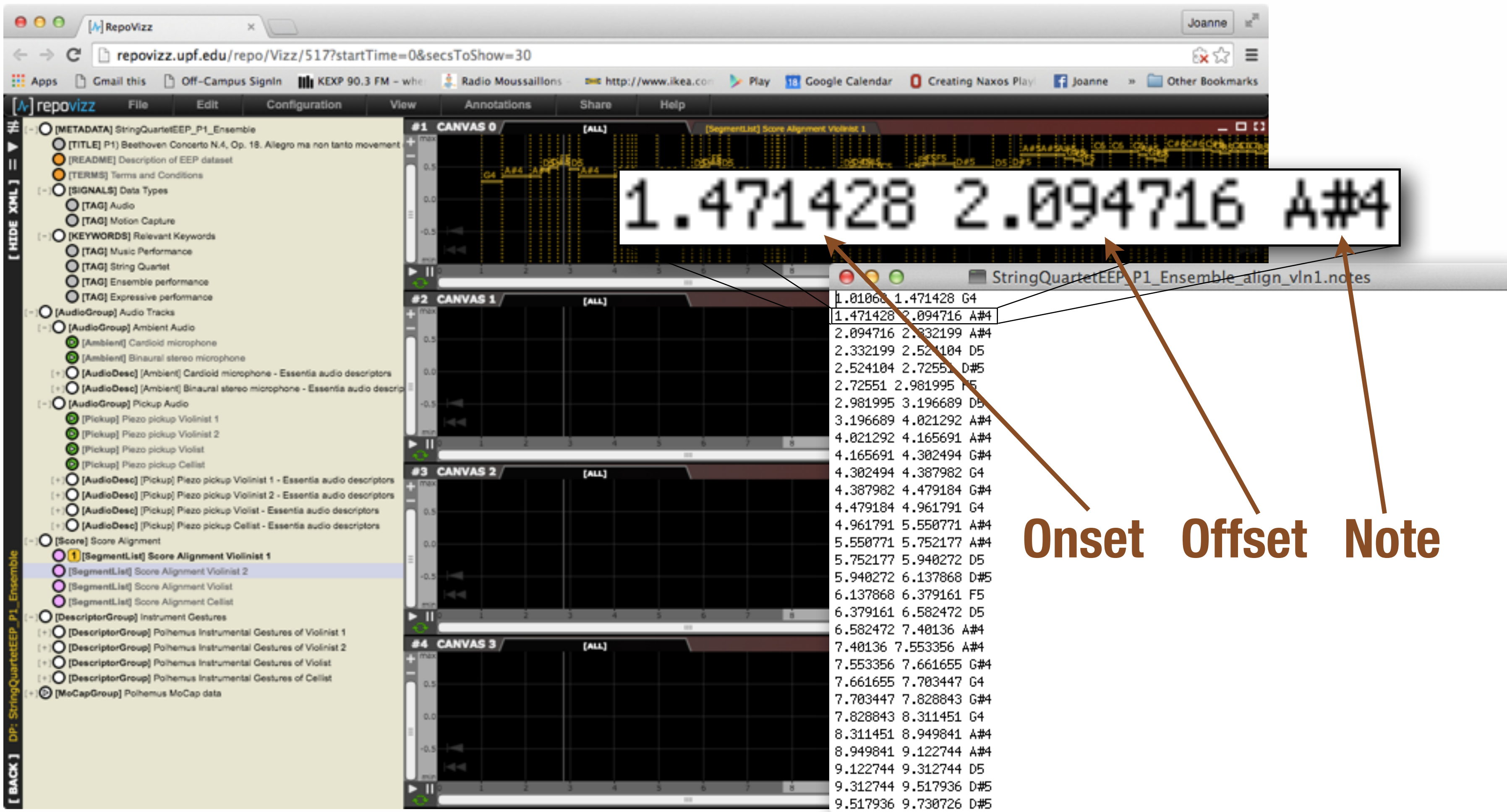
- **Ensemble Expressive Performance Dataset** (Marchini, Ramirez, Papiotis, and Maestre, 2014)



Current encoding formats

Note level

- **Ensemble Expressive Performance Dataset**
  (Marchini, Ramirez, Papiotis, and Maestre, 2014)



Current encoding formats

Note level

- **QMUL Singing Dataset**
  (Dai, Mauch, and Dixon, 2015)

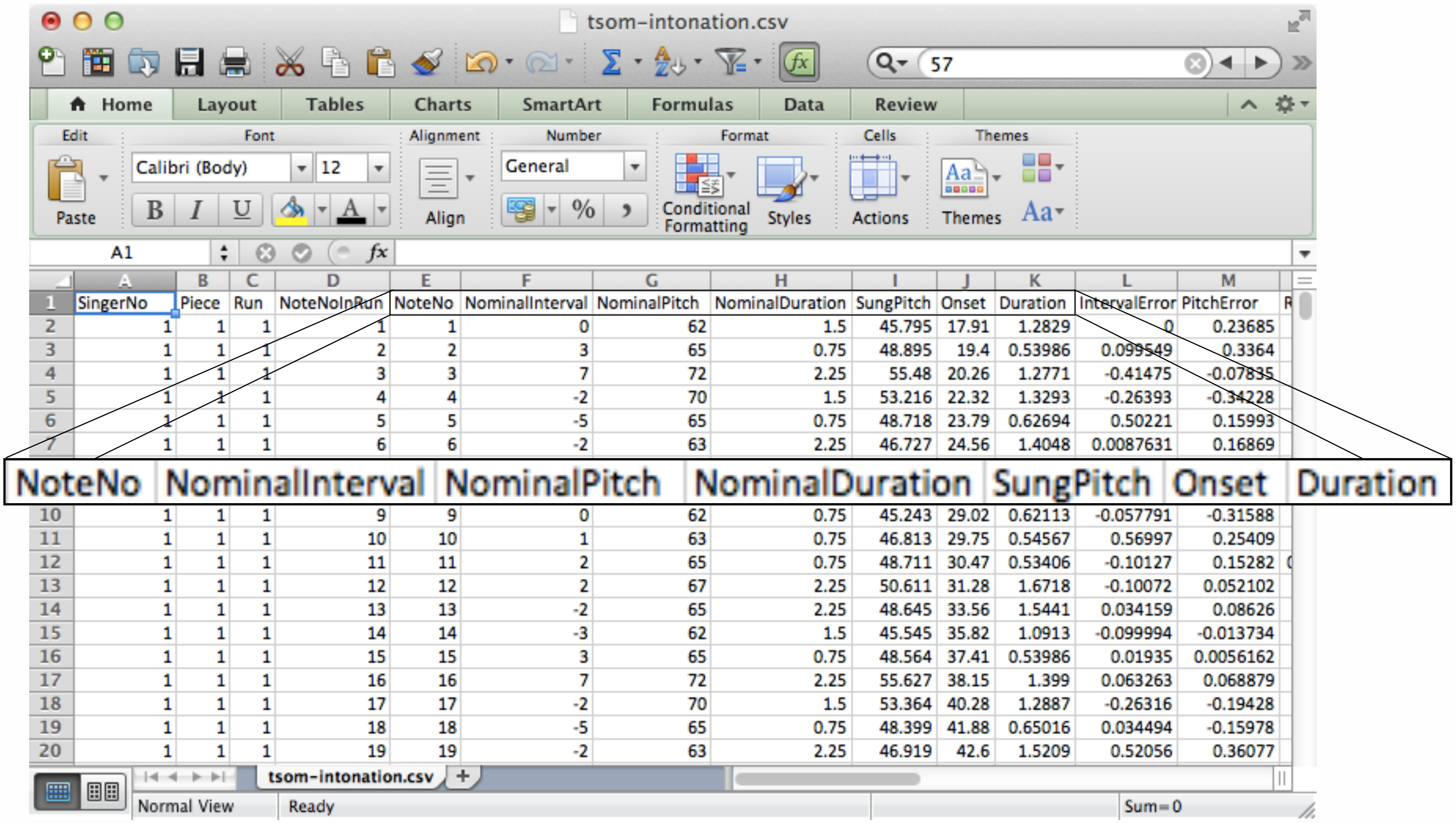
tsom-intonation.csv

SingerNo	Piece	Run	NoteNoInRun	NoteNo	NominalInterval	NominalPitch	NominalDuration	SungPitch	Onset	Duration	IntervalError	PitchError
1	1	1	1	1	0	62	1.5	45.795	17.91	1.2829	0	0.23685
1	1	1	2	2	3	65	0.75	48.895	19.4	0.53986	0.099549	0.3364
1	1	1	3	3	7	72	2.25	55.48	20.26	1.2771	-0.41475	-0.07835
1	1	1	4	4	-2	70	1.5	53.216	22.32	1.3293	-0.26393	-0.34228
1	1	1	5	5	-5	65	0.75	48.718	23.79	0.62694	0.50221	0.15993
1	1	1	6	6	-2	63	2.25	46.727	24.56	1.4048	0.0087631	0.16869
1	1	1	7	7	-1	62	1.5	45.532	26.9	0.91138	-0.19524	-0.026549
1	1	1	8	8	0	62	0.75	45.3	28.33	0.51664	-0.23154	-0.25809
1	1	1	9	9	0	62	0.75	45.243	29.02	0.62113	-0.057791	-0.31588
1	1	1	10	10	1	63	0.75	46.813	29.75	0.54567	0.56997	0.25409
1	1	1	11	11	2	65	0.75	48.711	30.47	0.53406	-0.10127	0.15282
1	1	1	12	12	2	67	2.25	50.611	31.28	1.6718	-0.10072	0.052102
1	1	1	13	13	-2	65	2.25	48.645	33.56	1.5441	0.034159	0.08626
1	1	1	14	14	-3	62	1.5	45.545	35.82	1.0913	-0.099994	-0.013734
1	1	1	15	15	3	65	0.75	48.564	37.41	0.53986	0.01935	0.0056162
1	1	1	16	16	7	72	2.25	55.627	38.15	1.399	0.063263	0.068879
1	1	1	17	17	-2	70	1.5	53.364	40.28	1.2887	-0.26316	-0.19428
1	1	1	18	18	-5	65	0.75	48.399	41.88	0.65016	0.034494	-0.15978
1	1	1	19	19	-2	63	2.25	46.919	42.6	1.5209	0.52056	0.36077

Current encoding formats

Note level

- **QMUL Singing Dataset**
  (Dai, Mauch, and Dixon, 2015)



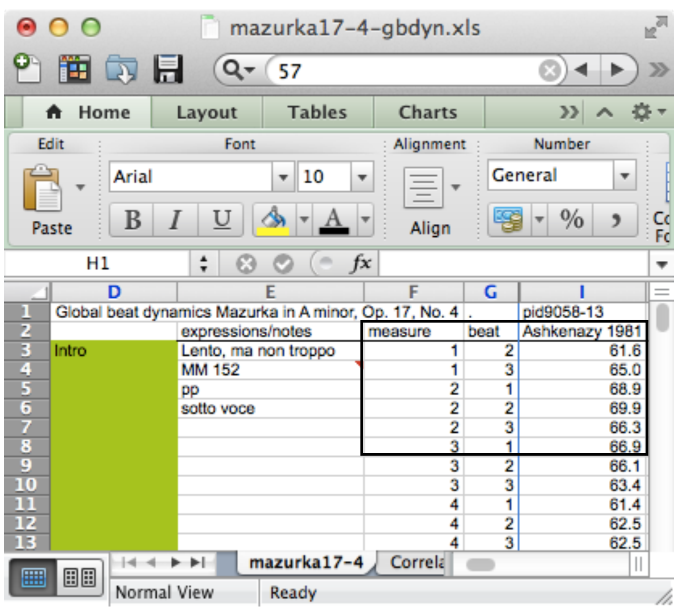
	SingerNo	Piece	Run	NoteNoInRun	NoteNo	NominalInterval	NominalPitch	NominalDuration	SungPitch	Onset	Duration	IntervalError	PitchError
2	1	1	1	1	1	0	62	1.5	45.795	17.91	1.2829	0	0.23685
3	1	1	1	2	2	3	65	0.75	48.895	19.4	0.53986	0.099549	0.3364
4	1	1	1	3	3	7	72	2.25	55.48	20.26	1.2771	-0.41475	-0.07835
5	1	1	1	4	4	-2	70	1.5	53.216	22.32	1.3293	-0.26393	-0.34228
6	1	1	1	5	5	-5	65	0.75	48.718	23.79	0.62694	0.50221	0.15993
7	1	1	1	6	6	-2	63	2.25	46.727	24.56	1.4048	0.0087631	0.16869
10	1	1	1	9	9	0	62	0.75	45.243	29.02	0.62113	-0.057791	-0.31588
11	1	1	1	10	10	1	63	0.75	46.813	29.75	0.54567	0.56997	0.25409
12	1	1	1	11	11	2	65	0.75	48.711	30.47	0.53406	-0.10127	0.15282
13	1	1	1	12	12	2	67	2.25	50.611	31.28	1.6718	-0.10072	0.052102
14	1	1	1	13	13	-2	65	2.25	48.645	33.56	1.5441	0.034159	0.08626
15	1	1	1	14	14	-3	62	1.5	45.545	35.82	1.0913	-0.099994	-0.013734
16	1	1	1	15	15	3	65	0.75	48.564	37.41	0.53986	0.01935	0.0056162
17	1	1	1	16	16	7	72	2.25	55.627	38.15	1.399	0.063263	0.068879
18	1	1	1	17	17	-2	70	1.5	53.364	40.28	1.2887	-0.26316	-0.19428
19	1	1	1	18	18	-5	65	0.75	48.399	41.88	0.65016	0.034494	-0.15978
20	1	1	1	19	19	-2	63	2.25	46.919	42.6	1.5209	0.52056	0.36077

Current encoding formats

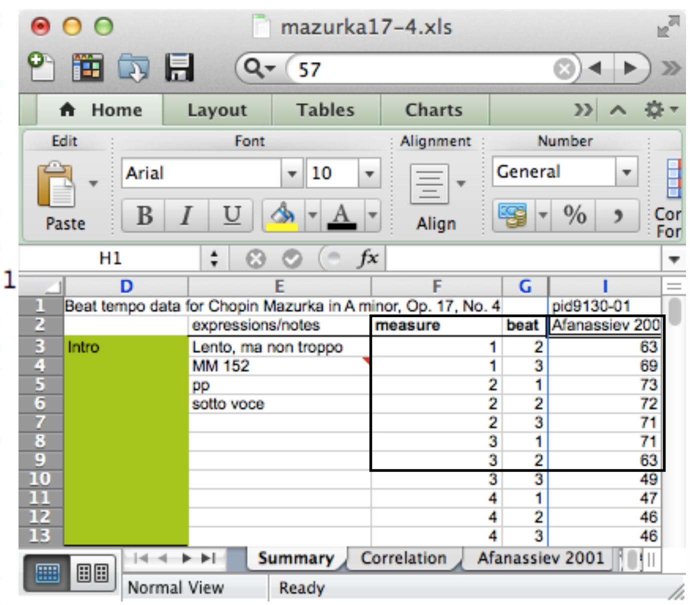
Beat level

- **Mazurka Dataset, CHARM**
  (Cook, 2007)

```
# C4DM/QMUL onset event detection algorithm output
# !!!pid:                pid9048-01
# !!!title:              Mazurka in F-sharp Minor, Op. 6
# !!!performer:          Chiu (1999)
# !!!analysis-channel:   Left
# !!!analysis-date:      Sat Mar 18 06:56:38 PST 2006
3.846   x1
4.728   x2
5.054   x3
5.274   x4
5.564   x5
5.739   x6
5.994   x7
6.098   x8
6.435   x9
6.598   x10
6.992   x11
7.143   x12
7.747   x13
8.444   x14
8.734   x15
8.908   x16
9.047   x17
9.140   x18
9.233   x19
9.338   x20
9.640   x21
9.732   x22
9.965   x23
10.127  x24
```



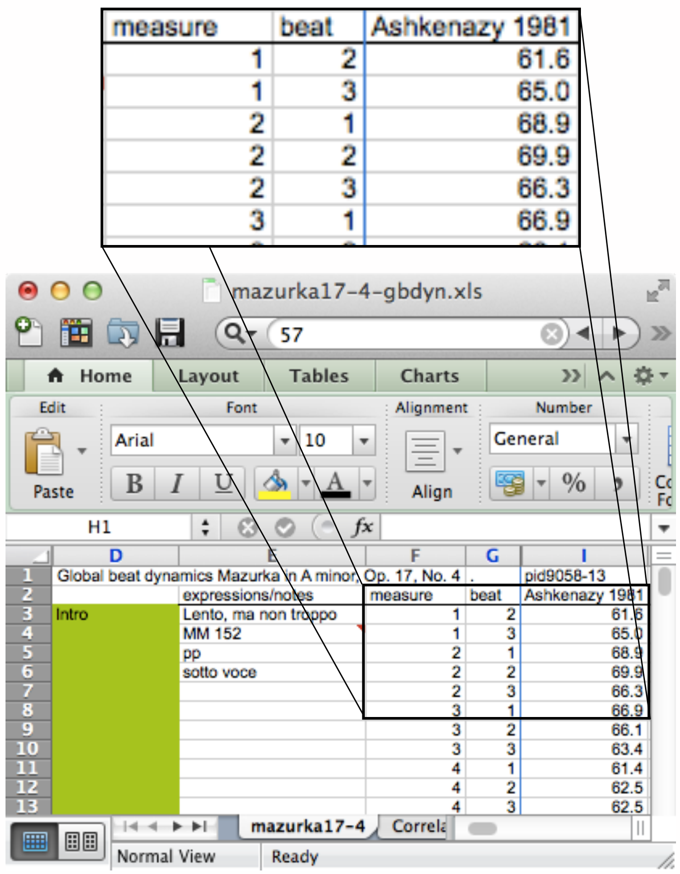

```
###PID:              pid9048-01
###FrameSize:        441 samples; 10 milliseconds
###FrameCount:       20324
###AnalysisDate:     Tue May  2 22:47:10 PDT 2006
###Comment:          Power contour for pid9048-01,
###Smoothing:        1
###Channel:          0 (left)
# time (seconds), dB
0       -99.8
0.01    -98.8
0.02    -99.2
0.03    -98.9
0.04    -99.8
0.05    -99.3
0.06    -99
0.07    -98.9
0.08    -99.9
0.09    -100.1
0.1     -99.1
0.11    -99.6
0.12    -98.7
0.13    -99
0.14    -99.9
0.15    -99.2
0.16    -99.2
0.17    -98.4
0.18    -99.1
0.19    -99.5
0.2     -99.9
```

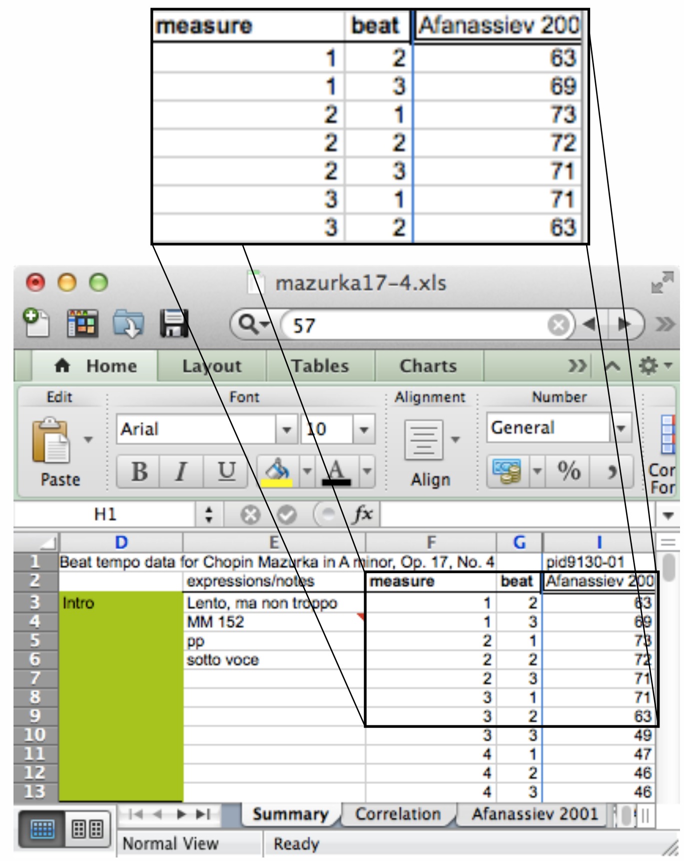


Current encoding formats

Beat level

- **Mazurka Dataset, CHARM**
  (Cook, 2007)





Current encoding formats

MIDI and MusicXML

- **MIDI**
  - *generative* rather than *descriptive*
  - available descriptors: micro-tuning, velocity, attack and release time, brightness
- **MusicXML**
  - extensions proposed by Hirata, Noike, and Katayose (2003) to incorporate note-level timing and dynamics information have not been implemented

Current encoding formats

Humdrum and MEI

- **Humdrum**
  - no specific guidelines for performance data in the standard Humdrum representation, but easily customizable (e.g., Craig Sapp's **time (tempo) and **idyn (attack velocities) spines)
- **MEI**
  - timing and temperament information can be encoded at either a broad scale with the <when> tag or at the note-level with the <note> tag, but specific performance-related attributes need to be defined

Introduction
Motivations and Background

1

Survey
Methods, Results, and Discussion

2

Encoding Format
Design and Implementation

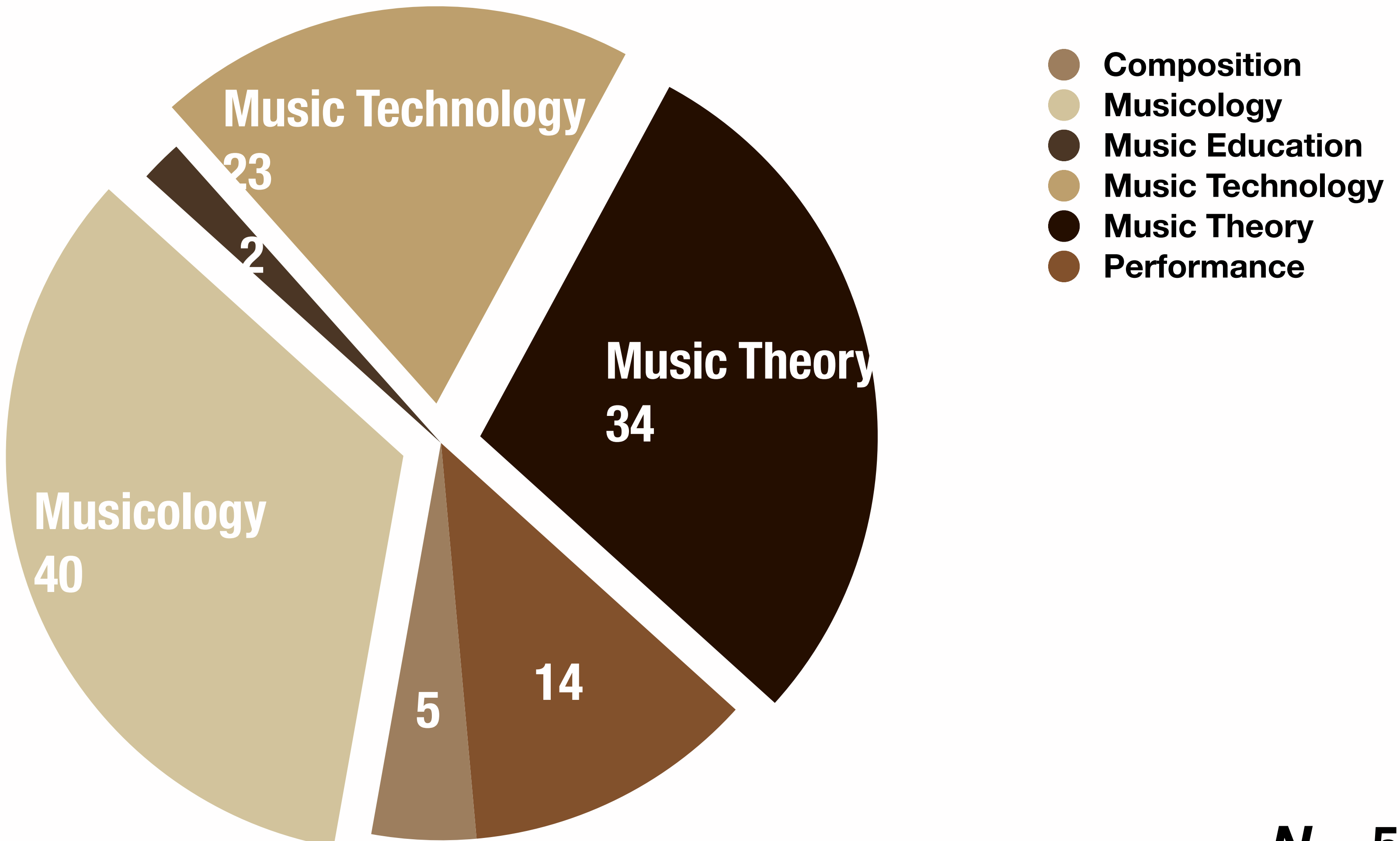
3

Conclusions
Future Directions and Summary

4

Demographics

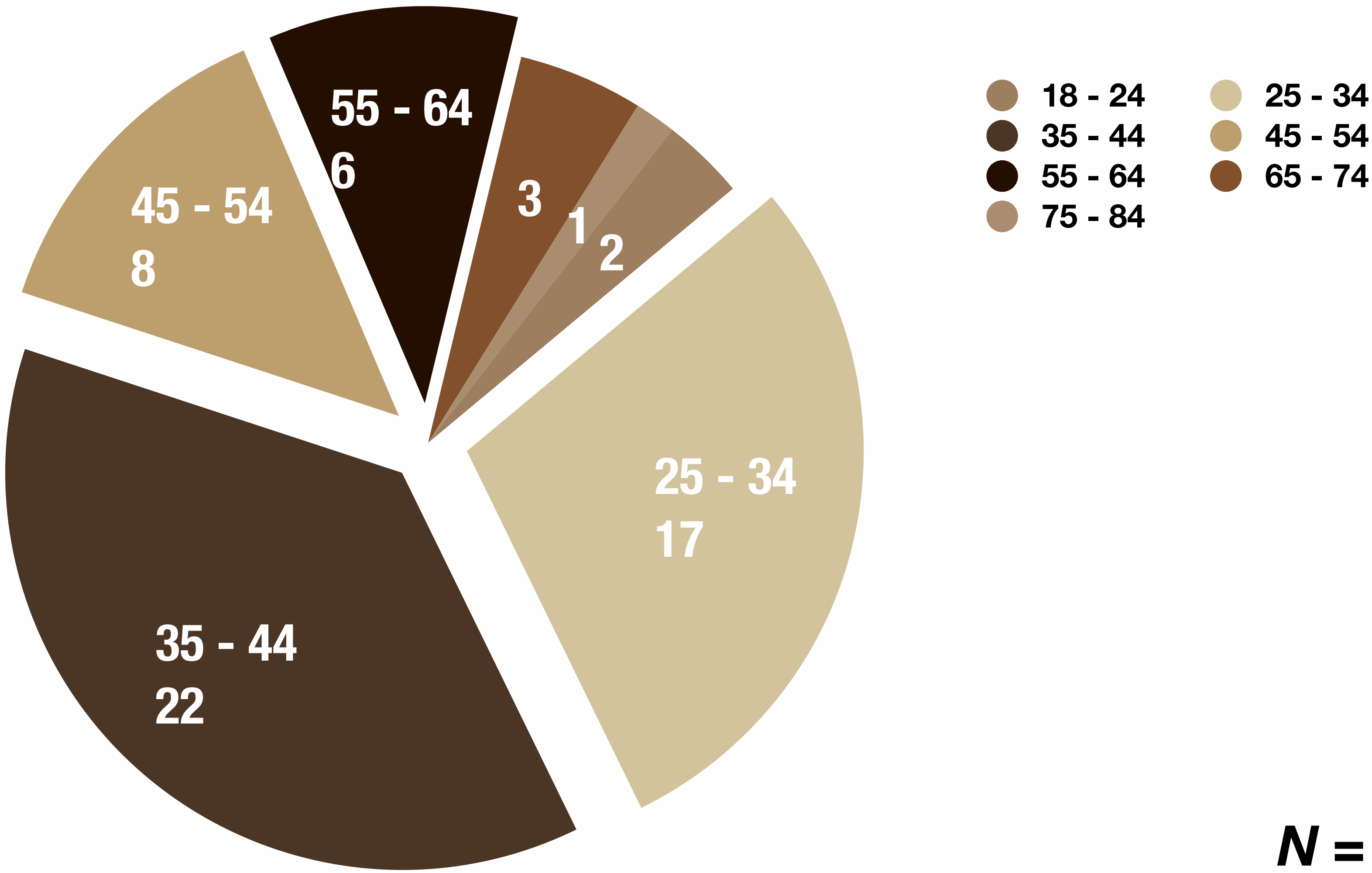
Music scholars from a range of sub-disciplines



N = 59

Demographics

Music scholars from a range of ages

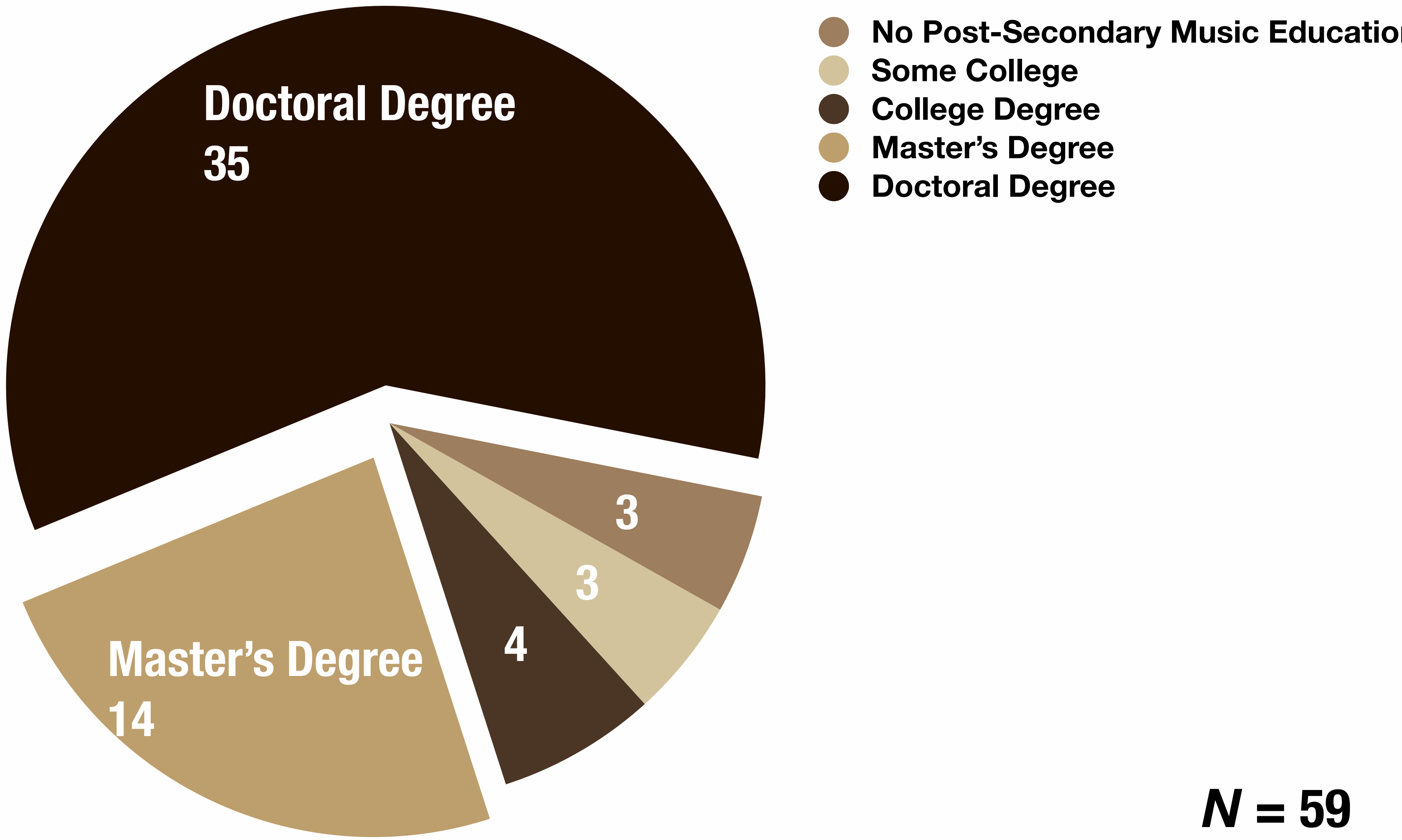


***N* = 59**

Not all respondents answered all questions

Demographics

Music scholars with a range of formal training

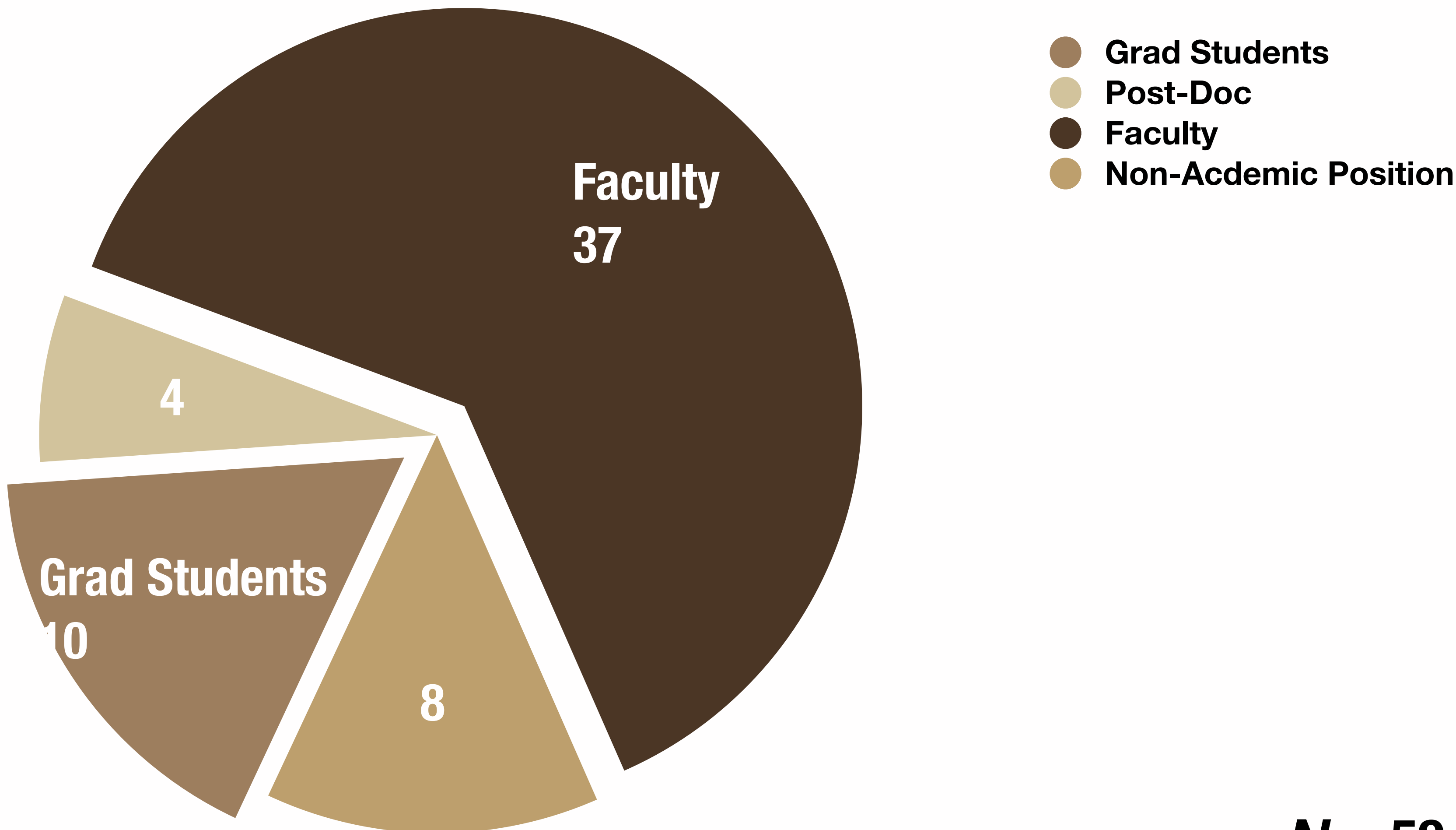


***N* = 59**

Not all respondents answered all questions

Demographics

Music scholars in a range of positions

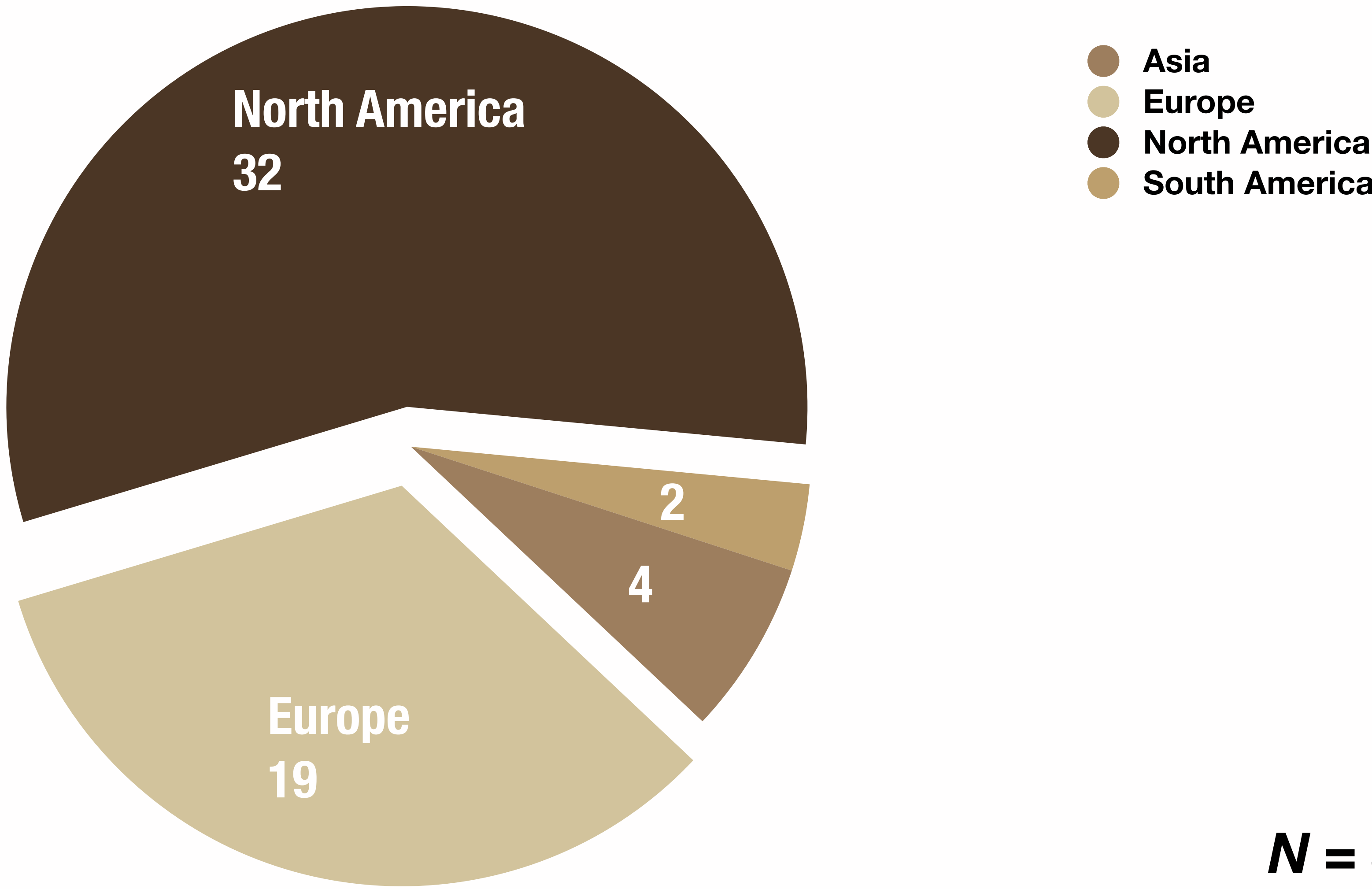


***N* = 59**

Not all respondents answered all questions

Demographics

Music scholars living in a range of locations

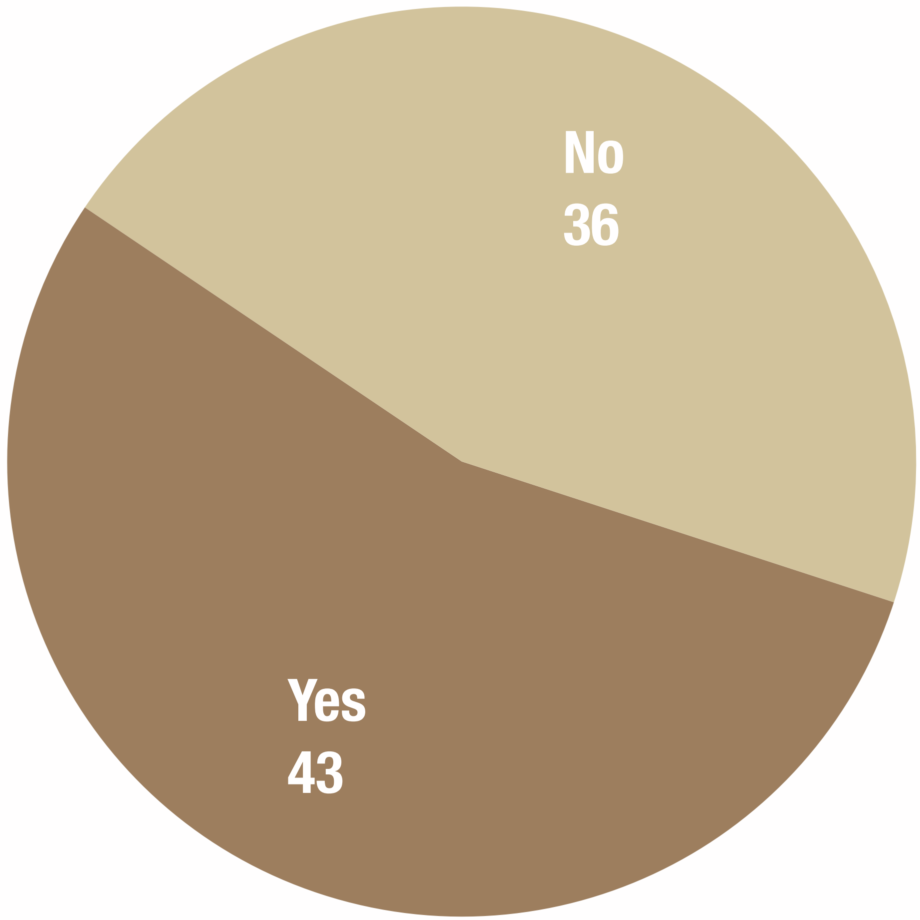


***N* = 59**

Not all respondents answered all questions

Experience

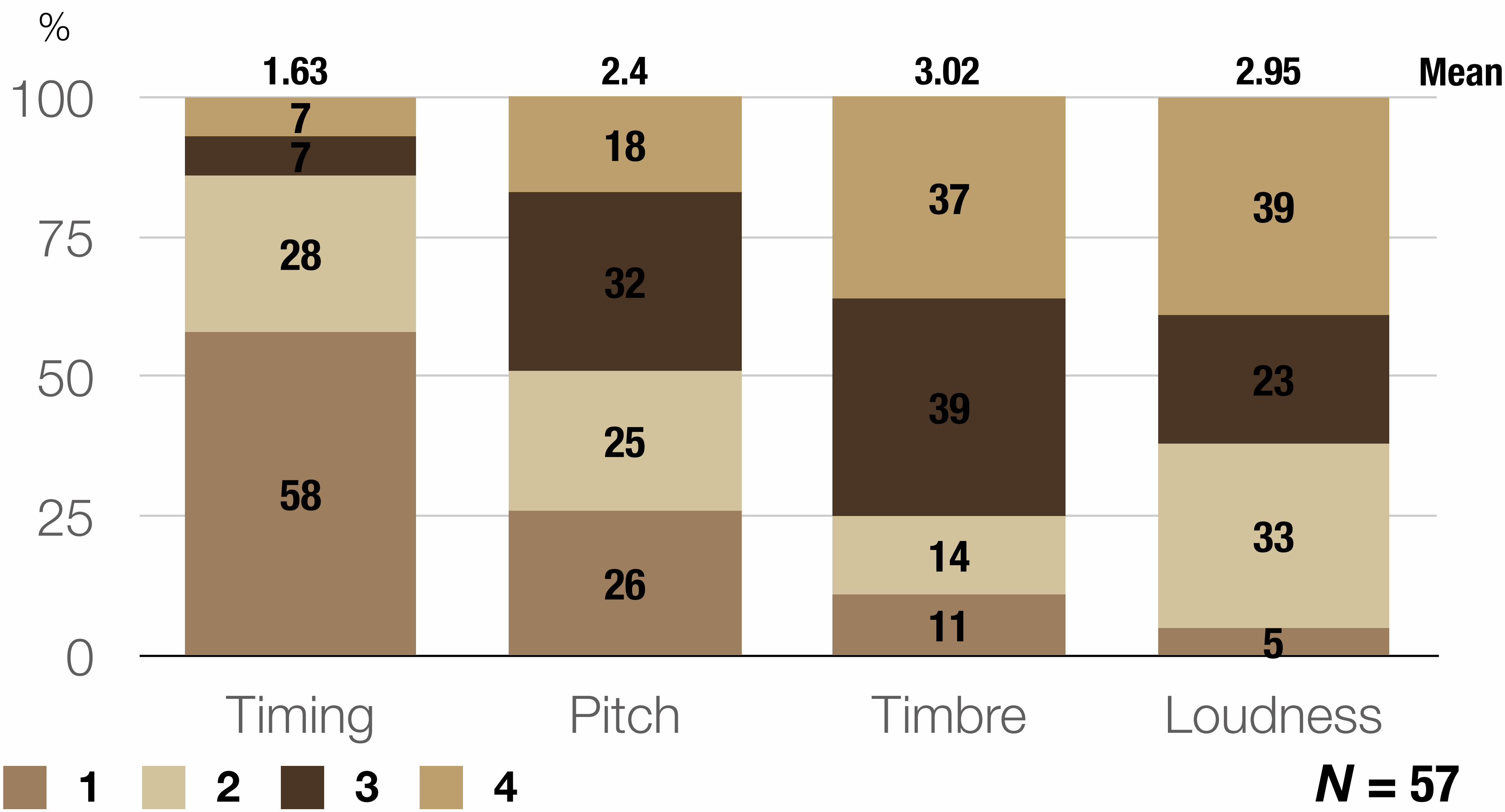
Music scholars reported a range of experience studying musical performance using computational tools



***N* = 79**

Musical Parameters

Music scholars reported varying interest in specific musical parameters that they would like to study/encode

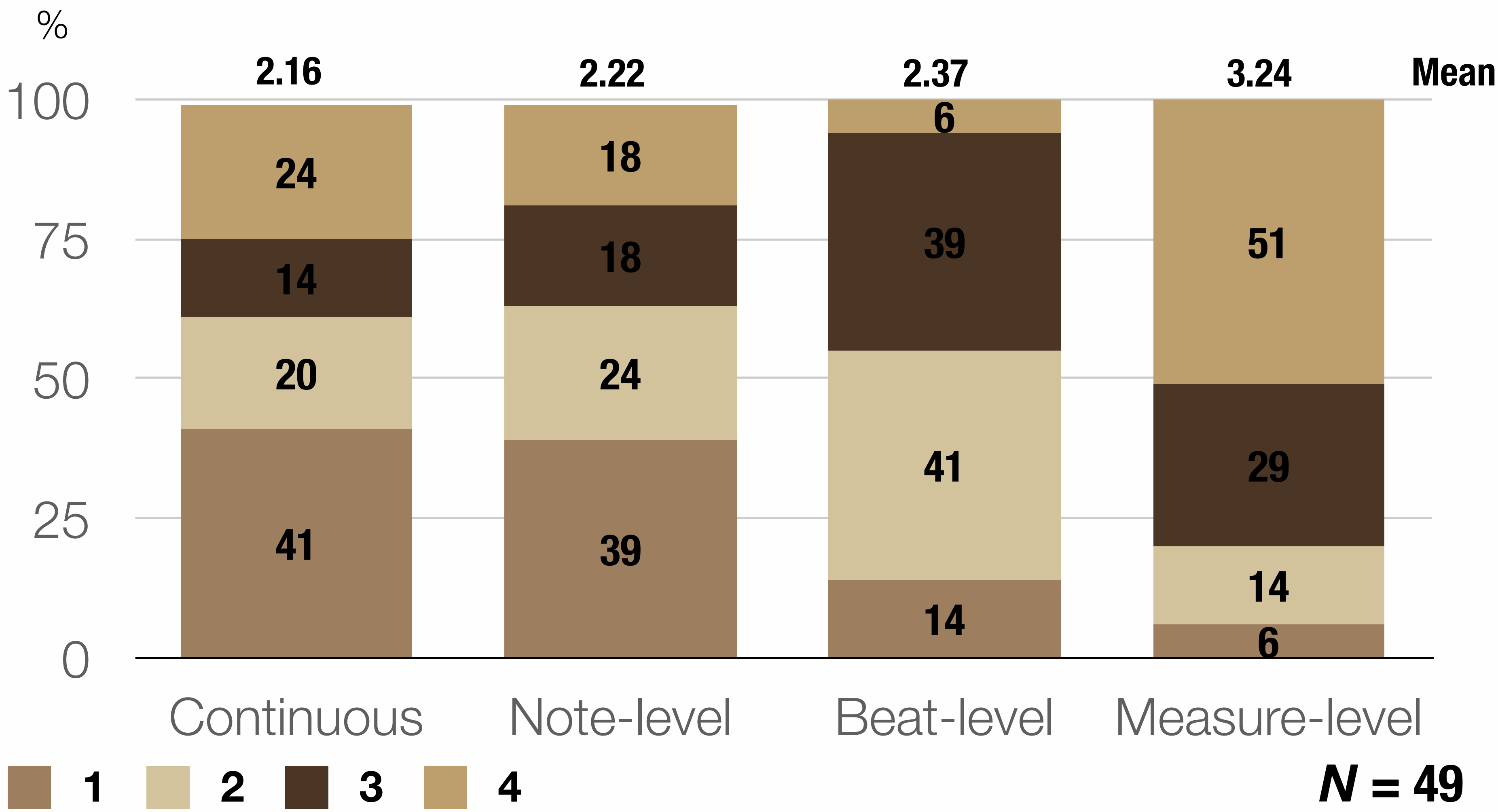


1 is most important

Not all respondents answered all questions

Musical Parameters

Music scholars reported varying interest in the level of description they would like to encode performance data at



1 is most important

Not all respondents answered all questions

Introduction
Motivations and Background
1

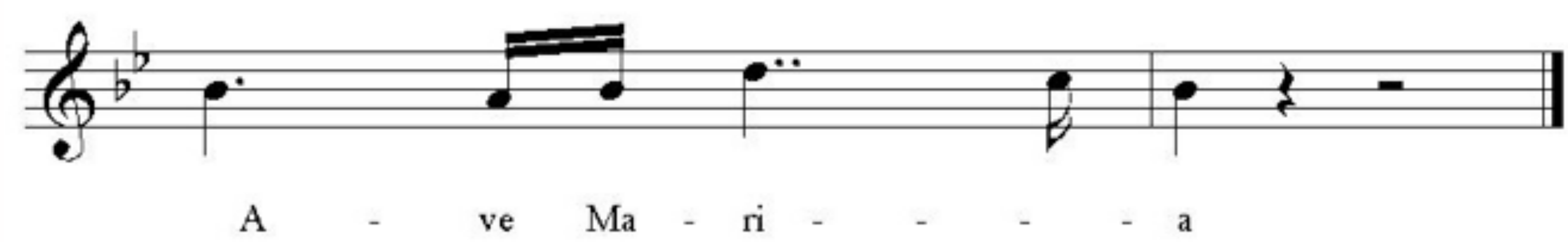
Survey
Methods, Results, and Discussion
2

Encoding Format
Design and Implementation
3

Conclusions
Future Directions and Summary
4

Design

Note-level, beat-level, measure-level, and continuous descriptors



Note-level	x		x	x	x		x	x
Beat-level	x	x			x	x		x
Measure-level	x							x
Continuous (general)	xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx							
Continuous (note-wise)	xxxxxxxx		xxx	xxx	xxxxxxxxx		xxx	xxxxxx

Encoding a vocal example

Note-level timing, loudness, pitch and timbre descriptors for the singing voice

Timing	***Loudness***
Inter-onset interval	Long-term loudness
Note duration	Shimmer (Mean)
Pitch	***Timbre***
Perceived pitch	Spectral flatness (Median)
Slope of F0 trajectory	Spectral slope (Median)
Curvature of F0 trajectory	Spectral flux (Median)
Vibrato depth	Spectral centroid (Median)
Vibrato rate	Harmonics-to-Noise ratio
Jitter (Mean)	

Encoding a vocal example

Note-level timing, loudness, pitch and timbre descriptors for the singing voice

Timing	*Loudness*
Inter-onset interval	Long-term loudness
Note duration	Shimmer (Mean)

Pitch	*Timbre*
Perceived pitch	Spectral flatness (Median)
Slope of F0 trajectory	Spectral slope (Median)
Curvature of F0 trajectory	Spectral flux (Median)
Vibrato depth	Spectral centroid (Median)
Vibrato rate	Harmonics-to-Noise ratio
Jitter (Mean)	

Encoding a vocal example

Note-level timing, loudness, pitch and timbre descriptors for the singing voice

Timing	Loudness
Inter-onset interval	Long-term loudness
Note duration	Shimmer (Mean)
Pitch	**Timbre**
Perceived pitch	Spectral flatness (Median)
Slope of F0 trajectory	Spectral slope (Median)
Curvature of F0 trajectory	Spectral flux (Median)
Vibrato depth	Spectral centroid (Median)
Vibrato rate	Harmonics-to-Noise ratio
Jitter (Mean)	

Encoding a vocal example

Note-level timing, loudness, pitch and timbre descriptors for the singing voice

Timing

Inter-onset interval

Note duration

Loudness

Long-term loudness

Shimmer (Mean)

Pitch

Perceived pitch

Slope of F0 trajectory

Curvature of F0 trajectory

Vibrato depth

Vibrato rate

Jitter (Mean)

Timbre

Spectral flatness (Median)

Spectral slope (Median)

Spectral flux (Median)

Spectral centroid (Median)

Harmonics-to-Noise ratio

Encoding a vocal example

Note-level timing, loudness, pitch and timbre descriptors for the singing voice

Timing	Loudness
Inter-onset interval	Long-term loudness
Note duration	Shimmer (Mean)
Pitch	**Timbre**
Perceived pitch	Spectral flatness (Median)
Slope of F0 trajectory	Spectral slope (Median)
Curvature of F0 trajectory	Spectral flux (Median)
Vibrato depth	Spectral centroid (Median)
Vibrato rate	Harmonics-to-Noise ratio
Jitter (Mean)	

Implementation

Note-level Humdrum example

Score

Performance

Timing Pitch Loudness Timbre

```
!!!COM: Schubert, Franz Peter
!!!CDT: 1797-1828
!!!OTL: Ellens Gesang III ("Ave Maria")
!!!OPS: Opus 52
!!!ONM: No. 6
!!!SCT: D 839
!!Only first 3 mm. of the vocal part
**beat	**kern	**silbe	**rtstart	**rtdur	**freq	**slopeF0	**curveF0	**vibdepth	**vibrate	**jitter	**loudness	**shimmer	**specflat	**specslope	**specflux	**speccent	**hnr
*	*clefG2	*M4/4	*u=msec	*u=msec	*clefG2	*	*	*	*	*	*	*	*	*	*	*	*
*	*k[b-e-]	*	*MM60	*MM60	*k[b-e-]	*	*	*	*	*	*	*	*	*	*	*	*
*M4/4	*M4/4	*	*	*	*M4/4	*	*	*	*	*	*	*	*	*	*	*	*
*MM60	*MM60	*	*	*	*MM60	*	*	*	*	*	*	*	*	*	*	*	*
=1	=1	=1	=1	=1	=1	=1	=1	=1	=1	=1	=1	=1	=1	=1	=1	=1	=1
1	4.b-	A-	0	2900.5	456	-0.10373	-0.15858	27	4.5	0.3	23.484	18.8	0.06444	-0.00518	0.00089	0.02867	0.10361
2.5	16a	ve	2900.5	417.56	433	-0.053198	-0.09806	35	4.8	0.3	20.344	17.5	0.0703	-0.00516	0.001	0.02989	0.10961
2.75	16b-	Ma-	3406	484.86	460	-0.00041444	-0.093974	16	6.2	0.3	23.818	17.2	0.06633	-0.00518	0.001	0.02873	0.1223
3	4..dd	ri-	3901	3380.7	581	-0.030237	-0.11501	10	4.1	0.1	32.989	11	0.05972	-0.00517	0.00055	0.02913	0.13322
4.75	16cc	.	7344.9	319.25	508	0.026659	-0.0016503	10	3.1	0.1	20.321	10.5	0.05956	-0.00519	0.00067	0.02836	0.12893
=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2
1	4b-	a	7814.6	2116.3	455	0.0057001	0.27152	20	4.3	0.2	21.27	8.7	0.05987	-0.00518	0.00078	0.02897	0.12201
2	4r
3	2r
==	==	==	==	==	==	==	==	==	==	==	==	==	==	==	==	==	==
```

Standard Humdrum notation: **beat – beat position **kern – note duration and name **silbe - lyrics

Timing Data
**rtstart - msec since start of first note
**rtdur – duration of note in msec

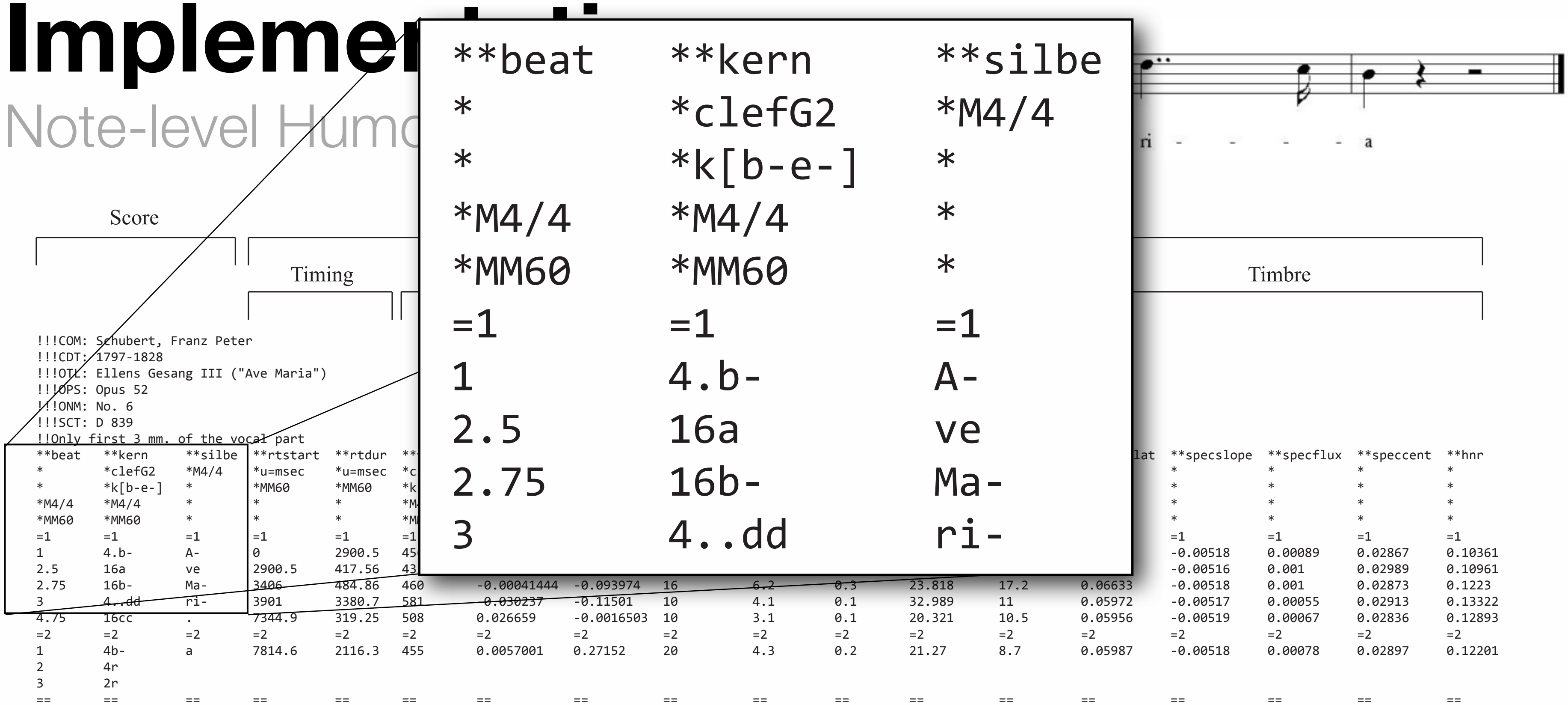
Pitch Data
**freq – perceived pitch
**slopeF0 – slope of fundamental frequency
**curveF0 – slope of fundamental frequency
**vibdepth – vibrato depth
**vibrate – vibrato rate
**jitter – mean frame-wise difference in F0

Loudness Data
**loudness – perceptual loudness in db
**shimmer - mean frame-wise difference in power/

Timbre Data
**specflat – mean spectral flatness
**specslope – mean spectral slope
**specflux – mean spectral flux
**speccent – mean spectral centroid
**hnr – mean harmonics to noise ratio

Implemen

Note-level Humd



Standard Humdrum notation: **beat – beat position **kern – note duration and name **silbe - lyrics

Timing Data
**rtstart - msec since start of first note
**rtdur – duration of note in msec

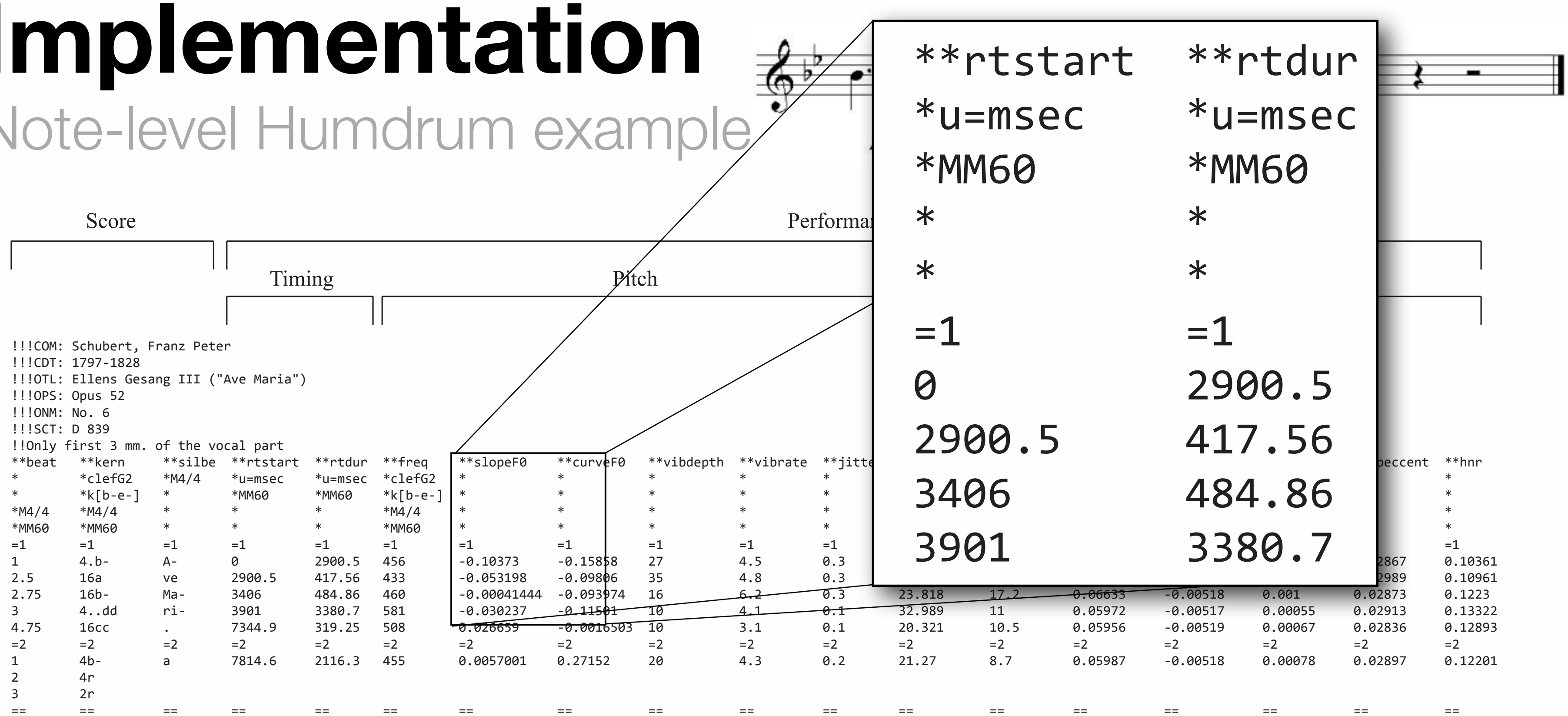
Loudness Data
**loudness – perceptual loudness in db
**shimmer - mean frame-wise difference in power/

Pitch Data
**freq – perceived pitch
**slopeF0 – slope of fundamental frequency
**curveF0 – slope of fundamental frequency
**vibdepth – vibrato depth
**vibrate – vibrato rate
**jitter – mean frame-wise difference in F0

Timbre Data
**specflat – mean spectral flatness
**specslope – mean spectral slope
**specflux – mean spectral flux
**speccent – mean spectral centroid
**hnr – mean harmonics to noise ratio

Implementation

Note-level Humdrum example



Standard Humdrum notation: **beat – beat position **kern – note duration and name **silbe - lyrics

Timing Data
**rtstart - msec since start of first note
**rtdur – duration of note in msec

Loudness Data
**loudness – perceptual loudness in db
**shimmer - mean frame-wise difference in power/

Pitch Data
**freq – perceived pitch
**slopeF0 – slope of fundamental frequency
**curveF0 – slope of fundamental frequency
**vibdepth – vibrato depth
**vibrate – vibrato rate
**jitter – mean frame-wise difference in F0

Timbre Data
**specflat – mean spectral flatness
**specslope – mean spectral slope
**specflux – mean spectral flux
**speccent – mean spectral centroid
**hnr – mean harmonics to noise ratio

Implementation

Note-level Humdrum example

Score

Performance

Timing | Pitch | Loudness | Timbre

```
!!!COM: Schubert, Franz Peter
!!!CDT: 1797-1828
!!!OTL: Ellens Gesang III ("Ave Maria")
!!!OPS: Opus 52
!!!ONM: No. 6
!!!SCT: D 839
!!Only first 3 mm. of the vocal part
**beat  **kern     **silbe  **rtstart  **rtdur   **freq     **slopeF0    **curveF0    **vibdepth  **vibrate  **jitter  **loudness  **shimmer  **specflat  **specslope  **specflux  **speccent  **hnr
*       *clefG2    *M4/4    *u=msec    *u=msec   *clefG2    *            *            *           *          *         *           *          *           *            *           *           *
*       *k[b-e-]   *        *MM60      *MM60     *k[b-e-]   *            *            *           *          *         *           *          *           *            *           *           *
*M4/4   *M4/4      *        *          *         *M4/4      *            *            *           *          *         *           *          *           *            *           *           *
*MM60   *MM60      *        *          *         *MM60      *            *            *           *          *         *           *          *           *            *           *           *
=1      =1         =1       =1         =1        =1         =1           =1           =1          =1         =1        =1          =1         =1          =1           =1          =1          =1
1       4.b-       A-       0          2900.5    456        -0.10373     -0.15858     27          4.5        0.3       23.484      18.8       0.06444     -0.00518     0.00089     0.02867     0.10361
2.5     16a        ve       2900.5     417.56    433        -0.053198    -0.09806     35          4.8        0.3       20.344      17.5       0.0703      -0.00516     0.001       0.02989     0.10961
2.75    16b-       Ma-      3406       484.86    460        -0.00041444  -0.093974    16          6.2        0.3       23.818      17.2       0.06633     -0.00518     0.001       0.02873     0.1223
3       4..dd      ri-      3901       3380.7    581        -0.030237    -0.11501     10          4.1        0.1       32.989      11         0.05972     -0.00517     0.00055     0.02913     0.13322
4.75    16cc       .        7344.9     319.25    508        0.026659     -0.0016503   10          3.1        0.1       20.321      10.5       0.05956     -0.00519     0.00067     0.02836     0.12893
=2      =2         =2       =2         =2        =2         =2                                                                                                                                      =2
1       4b-        a        7814.6     2116.3    455        0.00                                                                                                                                    0.12201
2       4r
3       2r
==      ==         ==       ==         ==        ==         ==                                                                                                                                      ==
```

**freq	**slopeF0	**curveF0	**vibdepth	**vibrate	**jitter
*clefG2	*	*	*	*	*
*k[b-e-]	*	*	*	*	*
*M4/4	*	*	*	*	*
*MM60	*	*	*	*	*
=1	=1	=1	=1	=1	=1
456	-0.10373	-0.15858	27	4.5	0.3
433	-0.053198	-0.09806	35	4.8	0.3
460	-0.00041444	-0.093974	16	6.2	0.3
581	-0.030237	-0.11501	10	4.1	0.1

Standard Humdrum notation: **bea

Timing Data

**rtstart - msec since start of first note

**rtdur – duration of note in msec

Pitch Data

**freq – perceived pitch

**slopeF0 – slope of fundamental frequency

**curveF0 – slope of fundamental frequency

**vibdepth – vibrato depth

**vibrate – vibrato rate

**jitter – mean frame-wise difference in F0

Timbre Data

**specflat – mean spectral flatness

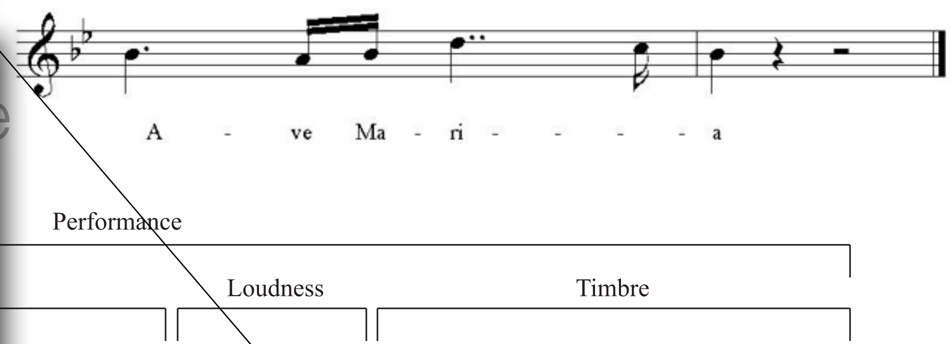
**specslope – mean spectral slope

**specflux – mean spectral flux

**speccent – mean spectral centroid

**hnr – mean harmonics to noise ratio

Implementation



**loudness	**shimmer
*	*
*	*
*	*
*	*
=1	=1
23.484	18.8
20.344	17.5
23.818	17.2
32.989	11

4.75	16cc	.	7344.9	319.25	508	0.026659	-0.0016503	10
=2	=2	=2	=2	=2	=2	=2	=2	=2
1	4b-	a	7814.6	2116.3	455	0.0057001	0.27152	20
2	4r							
3	2r							
==	==	==	==	==	==	==	==	==

**vibrate	**jitter	**loudness	**shimmer	**specflat	**specslope	**specflux	**speccent	**hnr
*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*
=1	=1	=1	=1	=1	=1	=1	=1	=1
4.5	0.3	23.484	18.8	0.06444	-0.00518	0.00089	0.02867	0.10361
4.8	0.3	20.344	17.5	0.0703	-0.00516	0.001	0.02989	0.10961
6.2	0.3	23.818	17.2	0.06633	-0.00518	0.001	0.02873	0.1223
4.1	0.1	32.989	11	0.05972	-0.00517	0.00055	0.02913	0.13322
3.1	0.1	20.321	10.5	0.05956	-0.00519	0.00067	0.02836	0.12893
=2	=2	=2	=2	=2	=2	=2	=2	=2
4.3	0.2	21.27	8.7	0.05987	-0.00518	0.00078	0.02897	0.12201
==	==	==	==	==	==	==	==	==

Standard Humdrum notation: **beat – beat position   **kern – note duration and name   **silbe - lyrics

Timing Data
**rtstart - msec since start of first note
**rtdur – duration of note in msec

Loudness Data
**loudness – perceptual loudness in db
**shimmer - mean frame-wise difference in power/

Pitch Data
**freq – perceived pitch
**slopeF0 – slope of fundamental frequency
**curveF0 – slope of fundamental frequency
**vibdepth – vibrato depth
**vibrate – vibrato rate
**jitter – mean frame-wise difference in F0

Timbre Data
**specflat – mean spectral flatness
**specslope – mean spectral slope
**specflux – mean spectral flux
**speccent – mean spectral centroid
**hnr – mean harmonics to noise ratio

Im

Not

Timbre

**specflat	**specslope	**specflux	**speccent	**hnr
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
=1	=1	=1	=1	=1
0.06444	-0.00518	0.00089	0.02867	0.10361
0.0703	-0.00516	0.001	0.02989	0.10961
0.06633	-0.00518	0.001	0.02873	0.1223
0.05972	-0.00517	0.00055	0.02913	0.13322

!!!COM:
!!!CDT:
!!!OTL:
!!!OPS:
!!!ONM:
!!!SCT:
!!Only f

**beat													**specflat	**specslope	**specflux	**speccent	**hnr
*													*	*	*	*	*
*													*	*	*	*	*
*M4/4													*	*	*	*	*
*MM60													*	*	*	*	*
=1													=1	=1	=1	=1	=1
1													0.06444	-0.00518	0.00089	0.02867	0.10361
2.5	16a	ve	2900.5	417.56	433	-0.053198	-0.09806	35	4.8	0.3	20.344	17.5	0.0703	-0.00516	0.001	0.02989	0.10961
2.75	16b-	Ma-	3406	484.86	460	-0.00041444	-0.093974	16	6.2	0.3	23.818	17.2	0.06633	-0.00518	0.001	0.02873	0.1223
3	4..dd	ri-	3901	3380.7	581	-0.030237	-0.11501	10	4.1	0.1	32.989	11	0.05972	-0.00517	0.00055	0.02913	0.13322
4.75	16cc	.	7344.9	319.25	508	0.026659	-0.0016503	10	3.1	0.1	20.321	10.5	0.05956	-0.00519	0.00067	0.02836	0.12893
=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2	=2
1	4b-	a	7814.6	2116.3	455	0.0057001	0.27152	20	4.3	0.2	21.27	8.7	0.05987	-0.00518	0.00078	0.02897	0.12201
2	4r																
3	2r																
==	==	==	==	==	==	==	==	==	==	==	==	==	==	==	==	==	==

Standard Humdrum notation: **beat – beat position **kern – note duration and name **silbe - lyrics

Timing Data
**rtstart - msec since start of first note
**rtdur – duration of note in msec

Pitch Data
**freq – perceived pitch
**slopeF0 – slope of fundamental frequency
**curveF0 – slope of fundamental frequency
**vibdepth – vibrato depth
**vibrate – vibrato rate
**jitter – mean frame-wise difference in F0

Loudness Data
**loudness – perceptual loudness in db
**shimmer - mean frame-wise difference in power/

Timbre Data
**specflat – mean spectral flatness
**specslope – mean spectral slope
**specflux – mean spectral flux
**speccent – mean spectral centroid
**hnr – mean harmonics to noise ratio

Implementation

Note-level MEI example

Timing Data
@rtstart - msec since start of first note
@rtdur – duration of note in msec

Loudness Data
@loudness – perceptual loudness in db
@shimmer - mean frame-wise difference in power

Pitch Data
@freq – perceived pitch
@slopeF0 – slope of fundamental frequency
@curveF0 – slope of fundamental frequency
@vibdepth – vibrato depth
@vibrate – vibrato rate
@jitter – mean frame-wise difference in F0

Timbre Data
@specflat – mean spectral flatness
@specslope – mean spectral slope
@specflux – mean spectral flux
@speccent – mean spectral centroid
@hnr – mean harmonics to noise ratio

```xml
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<mei>
    <meihead>
      <fileDesc>
            <titleStmt>
                  <title>Ellens Gesang III ("Ave Maria"), D 839</title>
                  <title type="subtitle">op. 52, no. 6</title>
                        <respStmt>
                              <persName role="creator">Schubert, Franz Peter</persName>
                        </respStmt>
            </titleStmt>
            <pubStmt>
                  <date>1797-1828</date>
            </pubStmt>
            <annot>Only first 3 mm. of the vocal part</annot>
      </fileDesc>
    </meihead>
```

```xml
<music>
   <body>
    <score>
        <scoreDef>
             meter.count="4"
             meter.unit="4"
             key.sig="2f" key.mode="major"
             clef.shape="G"
             clef.line="2"</scoreDef>
        <section>
             <staff n="1">
                  <measure n="1">
                       <note
                            xml:id="note_1"
                            tstamp="1"
                            dur="4"
                            dots="1"
                            oct="4"
                            pname="b"
                            accid.ges="f"
                            syl="A-"/>
                       <note
                            xml:id="note_2"
                            tstamp="2.5"
                            dur="16"
                            oct="4"
                            pname="a"
                            syl="ve"/>
                       <note
                            xml:id="note_3"
                            tstamp="2.75"
                            dur="16"
                            oct="4"
                            pname="b"
                            accid.ges="f"
                            syl="Ma-"/>
                       <note
                            xml:id="note_4"
                            tstamp="3"
                            dur="4"
                            dots="2"
                            oct="5"
                            pname="d"
                            syl="ri-"/>
                       <note
                            xml:id="note_5"
                            tstamp="4.75"
                            dur="16"
                            oct="5"
                            pname="c"/>
                  </measure>
                  <measure n="2">
                       <note
                            xml:id="note_6"
                            tstamp="4.75"
                            dur="4"
                            oct="4"
                            pname="b"
                            accid.ges="f"
                            syl="a"/>
                       <rest
                            tstamp="4.75"
                            dur="4"/>
                       <rest
                            tstamp="4.75"
                            dur="2"/>
                  </measure>
             </staff>
        </section>
    </score>
   </body>
```

```xml
<performance>
          <when xml:id="note_1"
                rtstart="00:00:00.0"
                rtdur="00:00:02.9005"
                freq="456"
                slopeF0="-0.10373"
                curveF0="-0.15858"
                vibdepth="27"
                vibrate="4.5"
                jitter="0.3"
                loudness="23.484"
                shimmer="18.8"
                specflat="0.06444"
                specslope="-.00518"
                specflux="0.00089"
                speccent="0.02867"
                hnr="0.10361" />
          <when xml:id="note_2"
                rtstart="00:00:02.9005"
                rtdur="00:00:00.41756"
                freq="433"
                slopeF0="-0.053198"
                curveF0="-0.09806"
                vibdepth="35"
                vibrate="4.8"
                jitter="0.3"
                loudness="20.344"
                shimmer="17.5"
                specflat="0.0703"
                specslope="-0.00516"
                specflux="0.001"
                speccent="0.02989"
                hnr="0.10961" />
          <when xml:id="note_3"
                rtstart="00:00:03.406"
                rtdur="00:00:00.48486"
                freq="460"
                slopeF0="-0.00041444"
                curveF0="-0.093974"
                vibdepth="16"
                vibrate="6.2"
                jitter="0.3"
                loudness="23.818"
                shimmer="17.2"
                specflat="0.06633"
                specslope="-0.00518"
                specflux="0.001"
                speccent="0.02873"
                hnr="0.1223" />
          <when xml:id="note_4"
                rtstart="00:00:03.901"
                rtdur="00:00:03.3807"
                freq="581"
                slopeF0="-0.030237"
                curveF0="-0.11501"
                vibdepth="10"
                vibrate="4.1"
                jitter="0.1"
                loudness="32.989"
                shimmer="11"
                specflat="0.05972"
                specslope="-0.000517"
                specflux="0.00055"
                speccent="0.02913"
                hnr="0.13322" />
          <when xml:id="note_5"
                rtstart="00:00:07.3449"
                rtdur="00:00:00.31925"
                freq="508"
                slopeF0="0.026659"
                curveF0="-0.0016503"
                vibdepth="10"
                vibrate="3.1"
                jitter="0.1"
                loudness="20.321"
                shimmer="10.5"
                specflat="0.05956"
                specslope="-0.00519"
                specflux="0.00067"
                speccent="0.02836"
                hnr="0.12893" />
          <when xml:id="note_6"
                rtstart="00:00:07.8146"
                rtdur="00:00:02.1163"
                freq="455"
                slopeF0="0.0057001"
                curveF0="0.27152"
                vibdepth="20"
                vibrate="4.3"
                jitter="0.2"
                loudness="21.27"
                shimmer="8.7"
                specflat="0.05987"
                specslope="-0.00518"
                specflux="0.00078"
                speccent="0.02897"
                hnr="0.12201" />
     </performance>
</music>
</mei>
```

Implementation

Note-level MEI example

Timing Data
@rtstart - msec since start of first note
@rtdur – duration of note in msec

Loudness Data
@loudness – perceptual loudness in db
@shimmer - mean frame-wise difference in power

Pitch Data
@freq – perceived pitch
@slopeF0 – slope of fundamental frequency
@curve
@vibde
@vibra
@jitter -

Timbre
@specf
@specs
@specf
@specc
@hnr –


```
<?xml version="
<mei>
  <meihea
    <fil
      </pubStmt>
      <annot>Only first 3 mm. of the vocal part</annot>
    </fileDesc>
  </meihead>
```

```
<music>
  <body>
    <score>
<scoreDef>
        meter.count="4"
        meter.unit="4"
        key.sig="2f" key.mode="major"
        clef.shape="G"
        clef.line="2"</scoreDef>
<section>
        <staff n="1">
                <measure n="1">
                        <note
                                xml:id="note_1"
                                tstamp="1"
                                dur="4"
                                dots="1"
                                oct="4"
                                pname="b"
                                accid.ges="f"
                                syl="A-"/>
                        <note
                                xml:id="note_2"
                                tstamp="2.5"
                                dur="16"
                                oct="4"
                                pname="a"
                                syl="ve"/>
                        <note
                                xml:id="note_3"
                                tstamp="2.75"
                                dur="16"
                                oct="4"
                                pname="b"
                                accid.ges="f"
                                syl="Ma-"/>
                        <note
                                xml:id="note_4"
                                tstamp="3"
                                dur="4"
                                dots="2"
                                oct="5"
                                pname="d"
                                syl="ri-"/>
                        <note
                                xml:id="note_5"
                                tstamp="4.75"
                                dur="16"
                                oct="5"
                                pname="c"/>
                </measure>
                <measure n="2">
                        <note
                                xml:id="note_6"
                                tstamp="4.75"
                                dur="4"
                                oct="4"
                                pname="b"
                                accid.ges="f"
                                syl="a"/>
                        <rest
                                tstamp="4.75"
                                dur="4"/>
                        <rest
                                tstamp="4.75"
                                dur="2"/>
                </measure>
        </section>
    </score>
  </body>
```

```
<performance>
        <when xml:id="note_1"
                rtstart="00:00:00.0"
                rtdur="00:00:02.9005"
                freq="456"
                slopeF0="-0.10373"
                curveF0="-0.15858"
                vibdepth="27"
                vibrate="4.5"
                jitter="0.3"
                loudness="23.484"
                shimmer="18.8"
                specflat="0.06444"
                specslope="-.00518"
                specflux="0.00089"
                speccent="0.02867"
                hnr="0.10361" />
        <when xml:id="note_2"
                rtstart="00:00:02.9005"
                rtdur="00:00:00.41756"
                freq="433"
                slopeF0="-0.053198"
                curveF0="-0.09806"
                vibdepth="35"
                vibrate="4.8"
                jitter="0.3"
                loudness="20.344"
                shimmer="17.5"
                specflat="0.0703"
                specslope="-0.00516"
                specflux="0.001"
                speccent="0.02989"
                hnr="0.10961" />
        <when xml:id="note_3"
                rtstart="00:00:03.406"
                rtdur="00:00:00.48486"
                freq="460"
                slopeF0="-0.00041444"
                curveF0="-0.093974"
                vibdepth="16"
                vibrate="6.2"
                jitter="0.3"
                loudness="23.818"
                shimmer="17.2"
                specflat="0.06633"
                specslope="-0.00518"
                specflux="0.001"
                speccent="0.02873"
                hnr="0.1223" />
        <when xml:id="note_4"
                rtstart="00:00:03.901"
                rtdur="00:00:03.3807"
                freq="581"
                slopeF0="-0.030237"
                curveF0="-0.11501"
                vibdepth="10"
                vibrate="4.1"
                jitter="0.1"
                loudness="32.989"
                shimmer="11"
                specflat="0.05972"
                specslope="-0.000517"
                specflux="0.00055"
                speccent="0.02913"
                hnr="0.13322" />
        <when xml:id="note_5"
                rtstart="00:00:07.3449"
                rtdur="00:00:00.31925"
                freq="508"
                slopeF0="0.026659"
                curveF0="-0.0016503"
                vibdepth="10"
                vibrate="3.1"
                jitter="0.1"
                loudness="20.321"
                shimmer="10.5"
                specflat="0.05956"
                specslope="-0.00519"
                specflux="0.00067"
                speccent="0.02836"
                hnr="0.12893" />
        <when xml:id="note_6"
                rtstart="00:00:07.8146"
                rtdur="00:00:02.1163"
                freq="455"
                slopeF0="0.0057001"
                curveF0="0.27152"
                vibdepth="20"
                vibrate="4.3"
                jitter="0.2"
                loudness="21.27"
                shimmer="8.7"
                specflat="0.05987"
                specslope="-0.00518"
                specflux="0.00078"
                speccent="0.02897"
                hnr="0.12201" />
        </performance>
</music>
</mei>
```

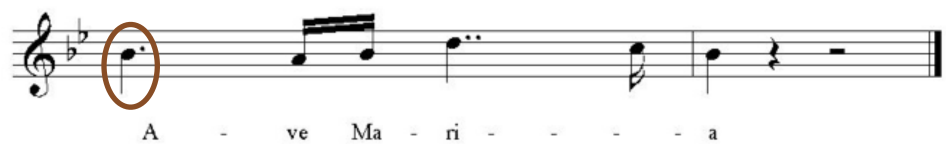


```
<scoreDef>
        meter.count="4"
        meter.unit="4"
        key.sig="2f" key.mode="major"
        clef.shape="G"
        clef.line="2"</scoreDef>
<section>
        <staff n="1">
                <measure n="1">
                        <note
                                xml:id="note_1"
                                tstamp="1"
                                dur="4"
                                dots="1"
                                oct="4"
                                pname="b"
                                accid.ges="f"
                                syl="A-"/>
```


Standard MEI note encoding

Implementation

Note-level MEI example



Timing Data
@rtstart - msec since start of first note
@rtdur – duration of note in msec

Loudness Data
@loudness – perceptual loudness in db
@shimmer - mean frame-wise difference in power

Pitch Data
@freq – perceived pitch
@slopeF0 – slope of fundamenta
@curveF0 – slope of fundamenta
@vibdepth – vibrato depth
@vibrate – vibrato rate
@jitter – mean frame-wise differer

Timbre Data
@specflat – mean spectral flatnes
@specslope – mean spectral slop
@specflux – mean spectral flux
@speccent – mean spectral centr
@hnr – mean harmonics to noise


```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<mei>
    <meihead>
      <fileDesc>
          <titleStmt>
              <title>Ellens Gesang III ("Ave Ma
              <title type="subtitle">op. 52, no
                  <respStmt>
                      <persName rol
                  </respStmt>
          </titleStmt>
          <pubStmt>
              <date>1797-1828</date>
          </pubStmt>
          <annot>Only first 3 mm. of the vocal part</annot>
      </fileDesc>
    </meihead>
```



```
<music>
    <body>
      <score>
        <scoreDef>
            meter.count="4"
            meter.unit="4"
            key.sig="2f" key.mode="major"
            clef.shape="G"
            clef.line="2"</scoreDef>
        <section>
            <staff n="1">
                <measure n="1">
                    <note
                        xml:id="note_1"
                        tstamp="1"
                        dur="4"
                        dots="1"
                        oct="4"
                        pname="b"
                        accid.ges="f"
                        syl="A-"/>
                    <note
                        xml:id="note_2"
                        tstamp="2.5"
        </section>
      </score>
    </body>
```



```
<performance>
        <when xml:id="note_1"
              rtstart="00:00:00.0"
              rtdur="00:00:02.9005"
              freq="456"
              slopeF0="-0.10373"
              curveF0="-0.15858"
              vibdepth="27"
              vibrate="4.5"
              jitter="0.3"
              loudness="23.484"
              shimmer="18.8"
              specflat="0.06444"
              specslope="-.00518"
              specflux="0.00089"
              speccent="0.02867"
              hnr="0.10361" />
```


Proposed performance encoding extension for MEI

Encoding a vocal example

Beat-level timing, loudness, and timbre descriptors

Timing	***Loudness***
T1: Inter-**beat** interval (Tempo)	L1: Long-term loudness
T2: **Beat** duration	~~L2: Shimmer (Mean)~~
Pitch	***Timbre***
P1: Interval size from previous note	Tb1: Mel Frequency Cepstral Coefficients 1–13 (Mean)
P2: Interval size from opening note	Tb2: Spectral flatness (Median)
P3: Slope of F0 trajectory	Tb3: Spectral slope (Median)
P4: Curvature of F0 trajectory	Tb4: Spectral flux (Median)
P5: Vibrato depth	Tb5: Spectral centroid (Median)
P6: Vibrato rate	~~Tb6: Harmonics-to-Noise ratio~~
P7: Jitter (Mean)	

Encoding a vocal example

Measure-level timing, loudness, and timbre descriptors

Timing	***Loudness***
T1: Inter-**down beat** interval	L1: Long-term loudness
T2: **Measure** duration	~~L2: Shimmer (Mean)~~
Pitch	***Timbre***
P1: Interval size from previous note	Tb1: Mel Frequency Cepstral Coefficients 1–13 (Mean)
P2: Interval size from opening note	Tb2: Spectral flatness (Median)
P3: Slope of F0 trajectory	Tb3: Spectral slope (Median)
P4: Curvature of F0 trajectory	Tb4: Spectral flux (Median)
P5: Vibrato depth	Tb5: Spectral centroid (Median)
P6: Vibrato rate	~~Tb6: Harmonics-to-Noise ratio~~
P7: Jitter (Mean)	

Beat- and Measure-Level

Potential problems

- **Simultaneously encoding more than one level can be problematic in Humdrum**

Beat- and Measure-Level

Potential problems

- **Simultaneously encoding more than one level can be problematic in Humdrum**

```
**beat   **kern     **silbe  **rtstart
*        *clefG2    *M4/4    *u=msec
*        *k[b-e-]   *        *MM60
*M4/4    *M4/4      *        *
*MM60    *MM60      *        *
=1       =1         =1       =1
1        4.b-       A-       0
2        .          .        2023.3
2.5      16a        ve       2900.5
2.75     16b-       Ma-      3406
3        4..dd      ri-      3901
4        .          .        6053.3
4.75     16cc       .        7344.9
=2       =2         =2       =2
1        4b-        a        7814.6
2        4r         .        9943.7
3        2r         .        11354.3
==       ==         ==       ==
```

Beat- and Measure-Level

Potential problems

- **Simultaneously encoding more than one level can be problematic in Humdrum**

```
**beat   **kern     **silbe  **rtstart
*        *clefG2    *M4/4    *u=msec
*        *k[b-e-]   *        *MM60
*M4/4    *M4/4      *        *
*MM60    *MM60      *        *
=1       =1         =1       =1
1        4.b-       A-       0
2        .          .        2023.3
2.5      16a        ve       2900.5
2.75     16b-       Ma-      3406
3        4..dd      ri-      3901
4        .          .        6053.3
4.75     16cc       .        7344.9
=2       =2         =2       =2
1        4b-        a        7814.6
2        4r         .        9943.7
3        2r         .        11354.3
==       ==         ==       ==
```

Beat- and Measure-Level

Potential problems

- **Simultaneously encoding more than one level can be problematic in Humdrum**

```
**beat   **kern    **silbe  **rtstart
*        *clefG2   *M4/4    *u=msec
*        *k[b-e-]  *        *MM60
*M4/4    *M4/4     *        *
*MM60    *MM60     *        *
=1       =1        =1       =1
1        4.b-      A-       0
2        .         .        2023.3
2.5      16a       ve       2900.5
2.75     16b-      Ma-      3406
3        4..dd     ri-      3901
4        .         .        6053.3
4.75     16cc      .        7344.9
=2       =2        =2       =2
1        4b-       a        7814.6
2        4r        .        9943.7
3        2r        .        11354.3
==       ==        ==       ==
```

```
**rtdur
*u=msec
*MM60
*
*
```

??

Beat- and Measure-Level

Potential problems

- **Simultaneously encoding more than one level can be problematic in Humdrum**

Beat- and Measure-Level

Potential problems

- **Simultaneously encoding more than one level can be problematic in Humdrum**
- **MEI is more flexible**

Beat- and Measure-Level

Potential problems

- **Simultaneously encoding more than one level can be problematic in Humdrum**
- **MEI is more flexible**

```
<section>
    <staff n="1">
        <measure n="1">
            <beat
                xml:id="measure_1_beat_1"
                tstamp="1"
                dur="4"/>
            <beat
                xml:id="measure_1_beat_2"
                tstamp="2"
                dur="4"/>
```

Beat- and Measure-Level

Potential problems

- **Simultaneously encoding more than one level can be problematic in Humdrum**
- **MEI is more flexible**

```
<section>
    <staff n="1">
        <measure n="1">
            <beat
                xml:id="measure_1_beat_1"
                tstamp="1"
                dur="4"/>
            <beat
                xml:id="measure_1_beat_2"
                tstamp="2"
                dur="4"/>
```

```
<performance>
    <when xml:id="measure_1_beat_1"
            rtstart="00:00:00.0"
            rtdur="00:00:02.4004"
            loudness="23.484"/>
    <when xml:id="measure_1_beat_2"
        rtstart="00:00:02.4004"
        rtdur="00:00:02.5065"
        loudness="24.427"/>
```

Continuous Data

- **Challenging in Humdrum because spines/ attributes only allow for single values rather than vectors of values**
- **Possible in MEI by using a <when> tag to specify exact time points**
- **An alternative is to use note-wise data summaries of continuous data (e.g., slope and curvature descriptors of F0)**
  - This also facilitates comparisons across notes of different lengths

Future Directions

Plans for the short-term

- **Implementation of note-wise descriptors**
  - Humdrum Extra tool
  - MEI module
- **Documentation for implementations**
  - Humdrum
  - MEI

Future Directions

Plans for the long-term

- **Addressing continuous data issues**
  - Specifically the ability to encoding continuous data rather than just summarizations
- **Addressing beat- and measure-level issues**
- **Developing tools to export data from formatted spreadsheets and to import data into Music21**

Summary

- **Need:** Demonstrated, through our survey, a need amongst music scholars for an encoding format that explicitly connects score information to performance data
- **Design:** Outlined encoding format for Humdrum and MEI
- **Implementation:** To come...

Thank you!

Thank you!

Questions/comments?

References

Cannam, C., C. Landone, M.B. Sandler, and J.P. Bello. 2006. The Sonic Visualiser: A Visualisation Platform for Semantic Descriptors from Musical Signals. In Proceedings of the International Conference on Music information Retrieval (ISMIR), 324–7.

Cook, N. 2007. Performance analysis and Chopin's mazurkas. Musicae Scientiae, 11(2), 183–207.

Cook, N. 2014. Beyond the Score: Music as Performance: Oxford University Press.

Dai, J., M. Mauch, and S. Dixon. 2015. Analysis of intonation trajectories in solo singing. In Proceedings of the International Society for Music Information Retrieval (ISMIR) conference. 42–6.

Devaney, J., M. Michael, and I. Fujinaga. 2012. A Study of Intonation in Three-Part Singing using the Automatic Music Performance Analysis and Comparison Toolkit (AMPACT) Proceedings of International Society for Music Information Retrieval Conference. 511–6.

Gingras, B., and S. McAdams. 2011. Improved score-performance matching using both structural and temporal information from MIDI recordings. Journal of New Music Research, 40(1): 43–57.

Hirata, K., K. Noike, and H. Katayose. 2003. A proposal for a performance data format. In Proceedings of the IJCAI-03 Workshop on Methods for Automatic Music Performance and their Applications in a Public Rendering Contest, 65–9.

Marchini, M., R. Ramirez, P. Papiotis, and E. Maestre. 2014. The sense of ensemble: a machine learning approach to expressive performance modelling in string quartets. Journal of New Music Research, 43 (3): 303–17.