Inter-/intra-performer similarity

Johanna Devaney
Assistant Professor of Music Theory and Cognition
School of Music
The Ohio State University
Introduction
Motivations.

A brief history
Quantitative approaches to performance analysis.

Inter-/Intra-singer similarity
Experiments with solo vocalists.

Conclusions
Summary and future directions.
Introduction
Similarity in performance

- Modeling style
  - style as self- or group-similarity
  - relationship between inter-performer similarity and intra-performer consistency
  - the need to sound spontaneous
    - Chaffin, Lemieux, Chen (2007)
Introduction

What do I mean by studying performance?

- Using (live) recorded performances
- Measuring performance parameters
  - timing
  - dynamics
  - tuning
  - timbre
- Assessing relationship between performance of various parameters and musical materials
Introduction
Motivations.

A brief history
Quantitative approaches to performance analysis.

Inter-/Intra-singer similarity
Experiments with solo vocalists.

Conclusions
Summary and future directions.
Quantitative Performance Analysis
A brief history

Pioneers
Binet and Courtier
Sears
Miller

1895–1930  1920–40s  1960s  1980s and 90s  1990s and 2000s
Quantitative Performance Analysis
A brief history

Pioneers
Binet and Courtier
Sears
Miller

1895–1930 1920–40s 1960s 1980s and 90s 1990s and 2000s

University of Iowa
Seashore and colleagues
Carl Seashore (1938) and colleagues studied timing, dynamics, intonation, and vibrato in pianists, violinists, and singers

- artistic performance conceived as deviations from the exact
Performance Scores
University of Iowa

Frequency curve
Loudness curve
Time

Seashore (1936)
How did Seashore model data?
Statistical methods used in Seashore’s lab

<table>
<thead>
<tr>
<th></th>
<th>Differences in extent</th>
<th>Differences in rate</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>% .1 step and less</td>
<td>% .2 step and less</td>
<td>Average</td>
</tr>
<tr>
<td>N cycles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baker</td>
<td>583</td>
<td>.07</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>Homer</td>
<td>207</td>
<td>.06</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>Kraft 1</td>
<td>168</td>
<td>.14</td>
<td>60</td>
<td>85</td>
</tr>
<tr>
<td>Kraft 2</td>
<td>201</td>
<td>.14</td>
<td>60</td>
<td>80</td>
</tr>
<tr>
<td>Marsh</td>
<td>428</td>
<td>.07</td>
<td>85</td>
<td>95</td>
</tr>
<tr>
<td>Seashore</td>
<td>303</td>
<td>.09</td>
<td>75</td>
<td>90</td>
</tr>
<tr>
<td>Stark</td>
<td>436</td>
<td>.07</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>Thompson</td>
<td>163</td>
<td>.09</td>
<td>80</td>
<td>95</td>
</tr>
<tr>
<td>Tibbett</td>
<td>260</td>
<td>.06</td>
<td>85</td>
<td>95</td>
</tr>
</tbody>
</table>

Seashore (1936)
Performance Scores

Digitizing the data
Quantitative Performance Analysis
A brief history

Pioneers
Binet and Courtier
Sears
Miller

1895–1930

1920–40s

1960s

1980s and 90s

1990s and 2000s

Ethnomusicology
Charles Seeger

University of Iowa
Seashore and colleagues
Quantitative Performance Analysis
A brief history

Pioneers
Binet and Courtier
Sears
Miller

Ethnomusicology
Charles Seeger

1895–1930
1920–40s
1960s
1980s and 90s
1990s and 2000s

University of Iowa
Seashore and colleagues

Piano
Gabrielsson
Todd
Clarke
Repp
Quantitative Performance Analysis

Popularity of the piano

- Large amount of solo repertoire
- Instrument’s percussive nature
- Feasibility of using specially equipped pianos (e.g., MIDI)
  - cannot study existing recordings
  - new recordings are typically done in a lab environment
How did these psychologists model data?
Statistical methods used in Repp’s piano studies

Averaging performances

Qualitative descriptions

Beethovenian  un-Beethovenian
Fast          slow
Expressive    inexpensive
Relaxed       tense
Superficial   deep
Cold          warm
Powerful      weak
Serious       playful
Pessimistic   optimistic
Smooth        rough
Spontaneous   deliberate
Consistent    variable
Coherent      incoherent
Sloppy        precise
Excessive     restrained
Rigid         flexible
Effortful     facile
Soft          hard
Realistic     idealistic
Usual         unusual

Repp (1990)
How did these psychologists model data?

Statistical methods used in Todd’s piano studies

Regression analysis

“the faster the louder, the slower the softer”

Todd (1992)
Quantitative Performance Analysis
A brief history

Pioneers
Binet and Courtier
Sears
Miller

Ethnomusicology
Charles Seeger

Other instruments
Fyk
Prame
Vurma

1895–1930
1920–40s
1960s
1980s and 90s
1990s and 2000s

University of Iowa
Seashore and colleagues

Piano
Gabrielsson
Todd
Clarke
Repp
Quantitative Performance Analysis
A brief history

### Pioneers
- Binet and Courtier
- Sears
- Miller

### Ethnomusicology
- Charles Seeger

### University of Iowa
- Seashore and colleagues

### Piano
- Gabrielsson
- Todd
- Clarke
- Repp

### Computational Models
- Friberg
- Mazola
- Widmer
- Sapp

### Other instruments
- Fyk
- Prame
- Vurma

**Timeline:**
- 1895–1930
- 1920–40s
- 1960s
- 1980s and 90s
- 1990s and 2000s
How do computer scientists model data?

Summary of statistical approaches used by Widmer et al.

**Case-based reasoning**

**Performance alphabets**

**Performance worms**

**Linear-basis functions**

---

Due to the huge efforts involved in manually measuring details of expressive timing from audio recordings, Repp’s analyses were limited to one particular piece. In our institute in Vienna, a large-scale project is currently being undertaken which aims at analysing truly large amounts of empirical performance data derived from recordings (Widmer et al., 2003). With the help of new computational methods that support the semi-automatic measurement of timing and dynamics from audio recordings, hundreds of recordings are being measured and characterised in terms of beat-level timing and global loudness changes.

7.1 Visualisation: performance trajectories

The resulting performance data – beat-level tempo and dynamics curves – can be represented in an integrated way as trajectories in a tempo–loudness space that show the joint development of tempo and dynamics over time (Langner & Goebl, 2003). Figure 1 shows a complete trajectory representing a performance of a Chopin Ballade by Artur Rubinstein. The line is produced by interpolating between the measured tempo and dynamics points, and smoothing the result with a Gaussian window to make the general trends visible. The degree of smoothing controls the amount of local variation that becomes visible.

A first intuitive analysis of high-level strategies characterising individual performances is facilitated by an interactive visualisation system called the Performance Worm (Dixon et al., 2002) that computes and visualises such performance trajectories via computer animation. But the trajectory representation also provides the basis for more detailed quantitative analysis, with data analysis (data mining) methods from the field of Artificial Intelligence.

Various avenues towards the characterisation of individual performance style are being followed, and we will briefly introduce some of these in the following subsections.

7.2 Characterisation: performance alphabets

The performance trajectories must first be converted into a form that is accessible to the automated data analysis machinery provided by data mining. To that end, the trajectories are cut into short segments of fixed length, e.g., two beats, which are then optionally subjected to various normalisation operations. The resulting segments can be grouped into classes of similar patterns via clustering. The centers of these clusters – the cluster prototypes – represent a set of typical elementary tempo–loudness patterns that can be used to approximately reconstruct a “full” trajectory (i.e., a complete performance). In that sense, they can be seen as a simple alphabet of performance, restricted to tempo and dynamics. Figure 2 displays such an alphabet computed from a set of Mozart sonata recordings by different artists. Such performance alphabets support a variety of quantitative analyses. A first useful step consists in the visualisation of the distribution of performance patterns over pianists, pieces, musical styles, etc. (Pampalk et al., 2003). That provides a very global view of aspects of personal style, such as “pianist A tends to use abrupt tempo turns combined with rather constant dynamics” or “pianist B combines a crescendo with a ritardando much more often than other pianists”. An example of such a visualisation can be found in (Widmer et al., 2003). An extensive study along these lines using Chopin performances by several famous pianists has recently revealed a number of characteristic performance strategies (Goebl et al., 2004).

---

**Figure 1.** Smoothed tempo–loudness trajectory representing a performance of Frédéric Chopin’s Ballade op.47 in Ab major by Artur Rubinstein. Horizontal axis: tempo in beats per minute (bpm); vertical axis: loudness in sone.

**Figure 2.** A “Mozart performance alphabet” (cluster prototypes) computed by segmentation, mean and variance normalization, and clustering, from performances of Mozart piano sonatas by six pianists (Daniel Barenboim, Roland Batik, Vladimir Horowitz, Maria João Pires, András Schiff, Mitsuko Uchida). To indicate directionality, dots mark the end points of segments. Shaded regions indicate the variance within a cluster.

---

**Tobudic and Widmer 2003**  
**Widmer and Goebl 2004**  
**Goebi, Pampalk, and Widmer 2004**  
**Grachten and Widmer 2012**
How do computer scientists model data?

Summary of statistical approaches used by Sapp

Dynascapes

Nearest-Neighbour

Sapp 2008
Piano data sets

What do they contain?

- **Vienna datasets (Bosendorfer)**
  - Magaloff performing the complete Chopin piano works
  - Batik performing 13 complete Mozart sonatas

- **Mazurka dataset (Commercial)**
  - 2926 recordings, between ~45–100 recordings per Chopin Mazurka – one recording per performer per era
  - Commercial recordings are a curated product
Introduction
Motivations.

A brief history
Quantitative approaches to performance analysis.

Inter-/Intra-singer similarity
Experiments with solo vocalists.

Conclusions
Summary and future directions.
Experiments with Singers

Overview

- Intonation in trained singers in the Western Art Music tradition
- Various aspect of the work was done in collaboration with Dan Ellis (Columbia), Ichiro Fujinaga (McGill), Michael Mandel (Ohio State), and Jon Wild (McGill)
Overview

Experiment design

- Musical Material
  - Schubert’s “Ave Maria”
    - 3x a cappella & 3x accompanied

- Singers
  - 6 non-professional singers: undergraduate vocal majors
  - 6 professional singers: possess at least one graduate-level degree in voice performance

- Melodic semitones and whole tones analyzed

- Singers listened to and approved their own recordings

Devaney, Mandel, Ellis and Fujinaga (2011)
Devaney, Wild, and Fujinaga (2011)
Data Extraction
Using MIDI-audio alignment

Darker Lines represent notes
Lighter lines represent non-notes

Lyrics
Alignment
Analysis

Vibrato and Intonation Information

<table>
<thead>
<tr>
<th>Interval Size</th>
<th>-64</th>
<th>72</th>
<th>388</th>
<th>210</th>
<th>202</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Vibrato Rate | 5.5 | 4 | 7.5 | 5 | 10 | 5 | Hertz |

Vibrato Depth | 26.5 | 10 | 5 | 30 | 3 | 33 | Cents |

Loudness: Glasberg and Moore (2002)


Pitch: Gockel, Moore, and Carlyon (2001)

Slope/Curvature: Devaney, Mandel and Fujinaga (2011)

www.ampact.org

Devaney, Mandel, and Fujinaga (2012)
Data Analysis

Linear regression

- **Dependent variable**
  - interval size in cents

- **Independent variables**
  - direction
  - singer or level of experience
  - harmonic context
    - leading tone or not
  - accompaniment
    - versus a cappella
Commonality between performers

Observable trends

-General tuning trends
  - No strict adherence, on average smaller than equal temperament (more so for semitones than whole tones)
  - Ascending semitones were significantly larger on average than descending semitones

-Harmonic context
  - Non-pros exhibited a significant difference between semitones in leading tone and non-leading tone contexts
    - Semitones in a leading context were significantly smaller on average
Is there an effect of training?
Professionals versus non-professionals

- **Effect of training**
  - **Accompaniment**
    - Solo non-pros’ accompanied semitones were 3 cents larger on average than their *a cappella* semitones
  - **Consistency**
    - Pros were more consistent with one another
  - **Interval size**
    - Pros’ semitones were significantly larger on average (closer to equal temperament)
Incorporating Seashore data

Comparative analysis of Seashore and contemporary data

**Semitones**

<table>
<thead>
<tr>
<th></th>
<th>Seashore</th>
<th>Devaney</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N = 166$</td>
<td>$N = 2266$</td>
</tr>
<tr>
<td><strong>Ascending</strong></td>
<td>105</td>
<td>99</td>
</tr>
<tr>
<td><strong>Descending</strong></td>
<td>100</td>
<td>96</td>
</tr>
</tbody>
</table>

**Whole tones**

<table>
<thead>
<tr>
<th></th>
<th>Seashore</th>
<th>Devaney</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N = 250$</td>
<td>$N = 2266$</td>
</tr>
<tr>
<td><strong>Ascending</strong></td>
<td>205</td>
<td>198</td>
</tr>
<tr>
<td><strong>Descending</strong></td>
<td>200</td>
<td>192</td>
</tr>
</tbody>
</table>
Singer Identity
Framing as a classification problem

Experiments

- Predicting singer identity within openings and closings using cross-validation
- Predicting singer identity of closing trained on opening

Support vector machine, with L1-regularization

- using the feature vectors for feature selection
## Singer Identity

Framing as a classification problem

<table>
<thead>
<tr>
<th>Pitch</th>
<th>Timing</th>
<th>Loudness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interval size</td>
<td>Inter-onset interval</td>
<td>Long-term loudness</td>
</tr>
<tr>
<td>Distance from opening note</td>
<td>Duration</td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curvature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vibrato extent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vibrato rate</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Cross-validation: A Cappella

Accuracy

<table>
<thead>
<tr>
<th></th>
<th>Non-professional</th>
<th>Professional</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Pitch</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Timing</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Loudness</td>
<td>70</td>
<td>70</td>
</tr>
</tbody>
</table>

Opening→Closing: A Cappella

Accuracy

<table>
<thead>
<tr>
<th></th>
<th>Non-professional</th>
<th>Professional</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Pitch</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Timing</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Loudness</td>
<td>70</td>
<td>70</td>
</tr>
</tbody>
</table>

Cross-validation: Accompanied

Accuracy

<table>
<thead>
<tr>
<th></th>
<th>Non-professional</th>
<th>Professional</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Pitch</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Timing</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Loudness</td>
<td>70</td>
<td>70</td>
</tr>
</tbody>
</table>

Opening→Closing: Accompanied

Accuracy

<table>
<thead>
<tr>
<th></th>
<th>Non-professional</th>
<th>Professional</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Pitch</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Timing</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Loudness</td>
<td>70</td>
<td>70</td>
</tr>
</tbody>
</table>
Introduction
Motivations.

A brief history
Quantitative approaches to performance analysis.

Inter-/Intra-singer similarity
Experiments with solo vocalists.

Conclusions
Summary and future directions.
Summary
Where we have been

- This talk has
  - provided a brief overview of the history of quantitative performance analysis with a particular focus on performance modeling
  - described the results of descriptive and predictive analysis of data from an experiment with twelve singers to explore inter- and intra-singer similarity
Future Work
Where might we be going?

- Different features
  - timbre

- More sophisticated musical models
  - looking at variance at particular points in the piece

- Categorical perception

- Integrating more qualitative information
  - performer intentionality
  - listener perception/reception
    - categorical perception of features – mid-level representation?
Acknowledgements

‣ School of Music and College of Arts and Sciences (OSU)
‣ Center for New Music and Audio Technologies (CNMAT)
‣ Distributed Digital Music Archives and Libraries (DDMAL)
‣ Centre for Research in Music Media and Technology (CIRMMT)
‣ Fonds de recherche sur la société et la culture (FQRSC)
‣ Social Sciences and Humanities Research Council of Canada (SSHRC)
‣ Advancing Interdisciplinary Research in Singing (AIRS)

Thank you!
References


