How do music experts and non-experts evaluate the vocal accuracy of operatic singing voices?

Larrouy-Maestri P., Nowak, M., & Morsomme D.

Voice Unit - Department of Psychology
University of Liège - Belgium
Evaluation of singing voice accuracy
Occasional singers

- **Music experts** (Larrouy-Maestri, Lévêque, Schön, Giovanni, & Morsomme, 2013)
  - High agreement between the judges
  - Correlation judges’ rating / pitch interval deviation

- **Non experts** (Larrouy-Maestri, Roig-Sanchis, & Morsomme, in prep)
  - Good intra- and inter-judge reliability
  - Correlation judges’ rating / pitch interval deviation

⇒ **Conclusions**
  - Shared definition of vocal accuracy
  - Importance of the interval deviation
  - Music experts seem better judges
Evaluation of singing voice accuracy
Operatic singers

- **Western operatic voices**
  - Complexity of the signal (e.g. Sundberg, 2013)
  - Parameters contributing to the beauty of the voice
    - (Ekholm et al., 1998; Garnier et al., 2007; Rothman et al., 1990)
  - Effect of these parameters on the perception
    - (e.g. Hutchins et al., 2012; Russo & Thompson, 2005; van Besouw et al., 2008; Vurma et al., 2010; Warrier & Zatorre, 2002)

- **Objectively out of tune**
  - Intervals (Vurma & Ross, 2006)
  - Melodic context (Larrouy-Maestri & Morsomme, 2012; Sundberg et al., 1996)
  - Whatever the melody performed (Larrouy-Maestri et al., in press)
  - Tolerance (Sundberg et al., 1996; Vurma & Ross, 2006)
  - Could serve the expressivity (Sundberg, La, & Himonides, in press)
Evaluation of operatic voices

- Methods
- Results
- Conclusions
Methods

Material

  - 50 professional singers
  - Two melodies: “Happy birthday” and “Romantic melody”
  - Two techniques: “Natural” and “Operatic singing technique”

- Selection of 14 performances
  - Female sung performances (from 245.42 Hz to 449.26 Hz, $M = 352.55$ Hz, $SD = 21.13$)
  - Last note long enough (from 1.13 s to 1.98 s, $M = 1.45$ s, $SD = 0.09$)
  - From 6 to 17 years of singing lessons ($M = 10.57$; $SD = 3.58$).
  - Practice: 14.07 hours/week on average
Methods
Performance and quality parameters

- **Performance parameters**
  - **Vocal accuracy** (AudioSculpt, OpenMusic, IRCAM, Paris, France)
  - F0 of the starting note (Hz)
  - Tempo (bpm)

- **Quality parameters (note 5)**
  - Energy distribution (2.4-5.4 kHz / total energy)
  - Vibrato rate (Hz)
  - Vibrato extent (cents)
# Methods

## Judges

<table>
<thead>
<tr>
<th></th>
<th>Experts</th>
<th>Non experts</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Gender</td>
<td>8 women</td>
<td>8 women</td>
</tr>
</tbody>
</table>
| Age           | From 26 to 73  
M = 45.68; SD = 11.16 | From 25 to 75  
M = 45.59; SD = 11.64 |
| Music expertise | From 15 to 55  
M = 35.77; SD = 10.74 | ___ |
| Practice      | Public performances  
M = 18.68 h/week | ___ |
| Audio         | ___      | OK          |
| MBEA (Peretz et al., 2003) | ___ | OK |
| Production task | ___ | OK |
Methods

Procedure

☐ Perceptual task
  ■ 14 selected melodies
  ■ Pairwise comparison paradigm: $N^*(N-1)/2 \Rightarrow 91$ pairs to compare
  ■ “Which one is the most in tune?”

☐ Judges’ evaluation of the performances
  ■ 1 point for “in tune”
  ■ 0 point for the other one
  ■ 0.5 for both when they are judged “equal”
  ■ Ranking of the performances for each judge (Kacha et al., 2005)

☐ Two times
  ■ 8-15 days in between
  ■ To observe the intra-judge reliability
Evaluation of operatic voices

• Methods
• Results
• Conclusions
Results and discussion

Intra-judge reliability

- **Variance between the test and the retest**

Experts

Non experts

- Better for music experts
Results and discussion
Inter-judge reliability

- **Correlations between the (reliable) judges**
  - Matrix of $\tau$ Kendall correlations
  - % of significant correlations ($p < .05$)

<table>
<thead>
<tr>
<th></th>
<th>Experts</th>
<th>Non experts</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>%</td>
<td>73.68</td>
<td>40</td>
</tr>
</tbody>
</table>

- **Better for music experts**
  - Non experts: different definitions of vocal accuracy
  - Experts: similar definition

- **What explains the music experts’ rating?**
Results and discussion
Definition of singing accuracy

- Spearman correlation judges’ rating / vocal accuracy
  - $r = .17; p = .56$
  - No direct relationship between subjective and objective evaluations

- Predictive model of vocal accuracy evaluation
  - All the performance and quality parameters
  - R2 coefficient: 78.8%

- Explanation of the judges’ rating
  - All the covariates appear
  - No main effect of one covariate
  - BUT always by means of an interacting effect with other covariate
Evaluation of operatic voices

• Methods
• Results
• Conclusions
Conclusions

Ability to evaluate singing voice

- **Occasional singers**
  - Music experts: ++
  - Non experts: +

- **Operatic singers**
  - Music experts: +
  - Non experts: -

ื่ Low variability (≠ Sundberg et al., 1996)
ี่ High consensus (≠ Ekholm et al., 1998; Garnier et al., 2007; Howes et al., 2004)
Conclusions
Definition of vocal accuracy

- **Occasional singers**
  - Interval deviation criterion

- **Operatic singers**
  - Interval deviation criterion
  - Complex combination of performance and quality parameters
  - Confirmation of the tolerance observed in intervallic and melodic contexts (Sundberg et al., 1996; Vurma & Ross, 2006)

- **Effect of expertise**
  - **Limited** for occasional singers’ performances
    - Enculturation and implicit learning are sufficient
  - **Important** for operatic singers’ performance
    - Musical training necessary to share the same definition
Conclusions

Perspectives

- **Pedagogy**
  - F0 variations + performance and quality parameters
  - Individual acoustical parameters cannot be observed separately
  - Less importance of the pitch

- **Research**
  - 21.2% of variance unexplained
    - Rhythm accuracy (Dalla Bella et al., 2007)
    - Vocal perturbation (Butte et al., 2009)
  - Synthesized material
    - To manipulate the performance and quality parameters
    - To precise their combination
    - To clarify the perception of vocal accuracy
  - Music expertise effect
The evaluation of operatic singing voices

THANK YOU FOR YOUR ATTENTION


References


- Larrouy-Maestri, P., Roig-Sanchis, V., & Morsomme, D. (2013, August 8th). Are we good judges even if we are not musicians. SMCP, Toronto, Canada.


References


