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

PEVOC
22/08/2013

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)

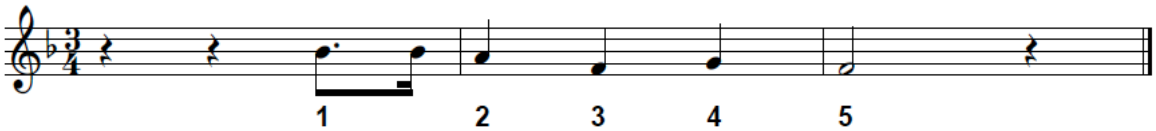
Evaluation of operatic voices

- Methods
- Results
- Conclusions

Methods

Material

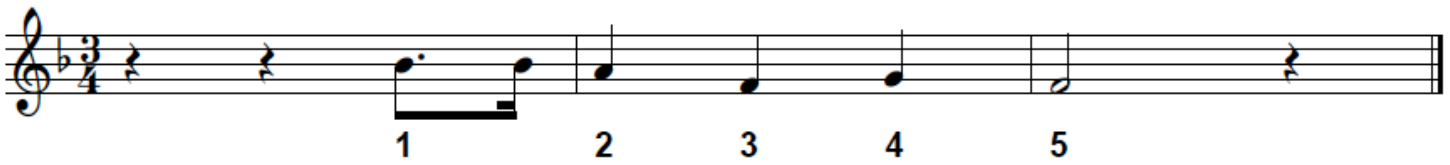
- Database: http://sldr.org/sldr000792/en
  - 50 professional singers
  - Two melodies: **"Happy birthday"** and "Romantic melody"
  - Two techniques: "Natural" and **"Operatic singing technique"**
- Selection of 14 performances
  - Female sung performances (from 245.42 Hz to 449.26 Hz, *M* = 352,55 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

	Experts	Non experts
n	22	22
Gender	8 women	8 women
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 ➔ 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

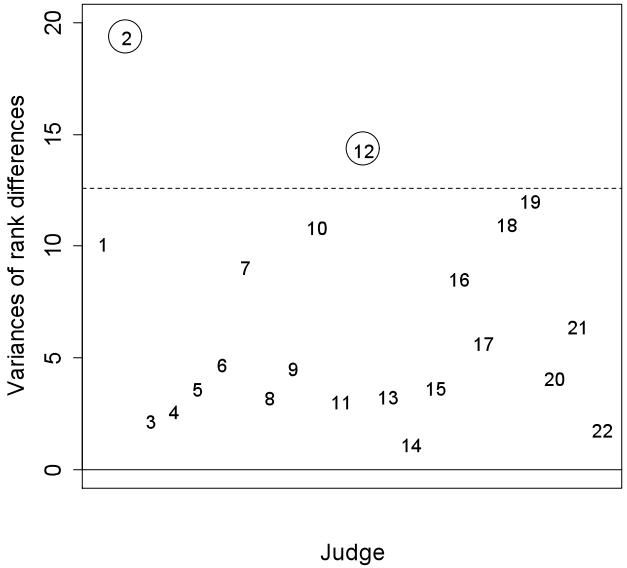
- Methods
- Results
- Conclusions

Results and discussion

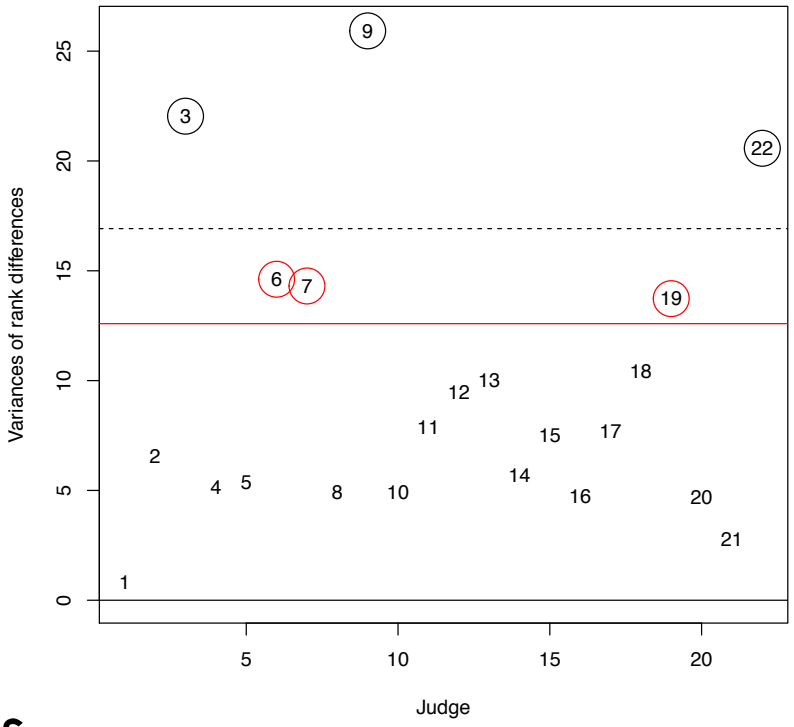
Intra-judge reliability

- 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 τ Kendall correlations
  - % of significant correlations ($p < .05$)

	Experts	Non experts
n	20	16
%	73.68	40

- 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

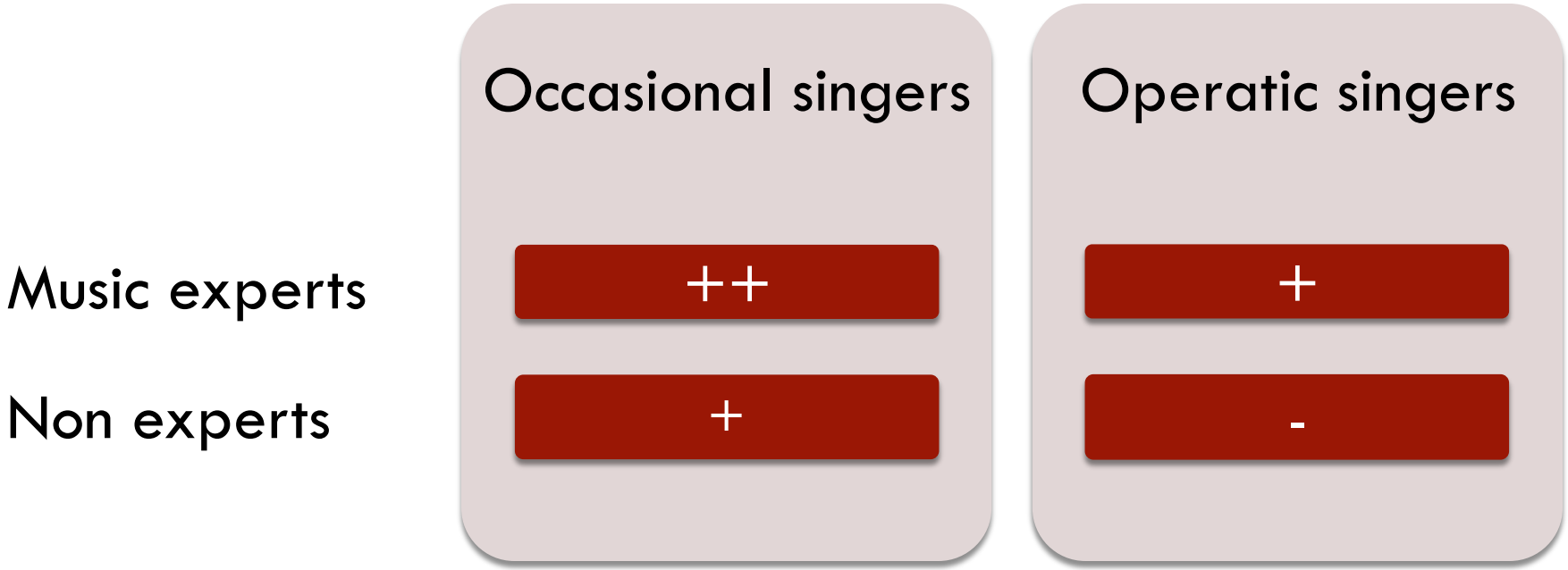
- Spearman correlation judges' rating / vocal accuracy
  - $r = .17$; $p = .56$
  - No direct relationship between subjective and objective evaluations
- Predictive model of vocal accuracy evaluation
  - All the performance and quality parameters
  - R2 coefficient: 78.8%
- Explanation of the judges' 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



➔ Low variability (≠ Sundberg et al., 1996)

➔ High consensus (≠ Ekholm et al., 1998; Garnier et al., 2007; Howes et al., 2004)

Conclusions

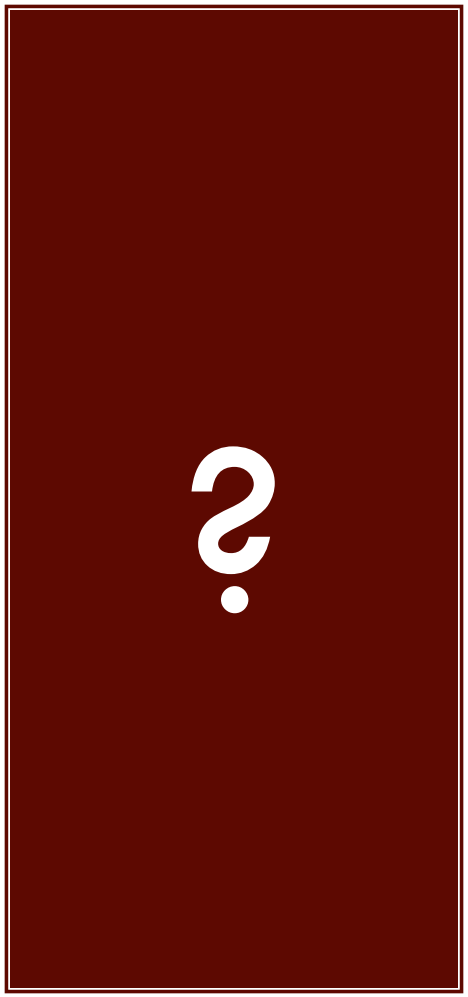
Definition of vocal accuracy

- 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



THANK YOU FOR YOUR ATTENTION

References

- Butte, C. J., Zhang, Y., Song, H., & Jiang, J. J. (2009). Perturbation and nonlinear dynamic analysis of different singing styles. *Journal of Voice, 23*, 647–652.
- Dalla Bella, S., Giguère, J.-F., & Peretz, I. (2007). Singing proficiency in the general population. The Journal of the Acoustical Society of America, 121(2), 1182-1189.
- Ekholm, E., Papagiannis, G. C., & Chagnon, F. P. (1998). Relating objective measurements to expert evaluation of voice quality in western classical singing: Critical perceptual parameters. *Journal of Voice, 12*, 182–196.
- Garnier, M., Henrich, N., Castellengo, M., Sotiropoulos, D., & Dubois, D. (2007). Characterisation of voice quality in Western lyrical singing: From teachers' judgements to acoustic descriptions. *Journal of Interdisciplinary Music Studies, 1*, 62–91.
- Howes, P., Callaghan, J., Davis, P., Kenny, D., & Thorpe, W. (2004). The relationship between measured vibrato characteristics and perception in Western operatic singing. *Journal of Voice, 18*, 216–230.
- Hutchins, S., Roquet, C., & Peretz, I. (2012). The Vocal Generosity Effect: How Bad Can Your Singing Be? Music Perception, 30(2), 147-159.
- Kacha, A., Grenez, F., & Schoentgen, J. (2005). Voice quality assessment by means of comparative judgments of speech tokens. Proceedings of *9th European Conference on Speech Communication and Technology, Interspeech* (pp. 1733-1736). Lisbon, Portugal.

References

- Larrouy-Maestri, P., Magis, D., & Morsomme, D. (in press). The effect of melody and technique on the singing voice accuracy of trained singers. *Logopedics, Phoniatrics, Vocology*.
- Larrouy-Maestri, P., & Morsomme, D. (in press). Criteria and tools for objectively analysing the vocal accuracy of a popular song. *Logopedics Phoniatrics, Vocology*.
- Larrouy-Maestri, P., Lévêque, Y., Schön, D., Giovanni, A., & Morsomme, D. (2013). The Evaluation of Singing Voice Accuracy: A Comparison Between Subjective and Objective Methods. Journal of voice, 27(2), 259.
- Larrouy-Maestri, P., Roig-Sanchis, V., & Morsomme, D. (2013, August 8th). Are we good judges even if we are not musicians. SMCP, Toronto, Canada.
- Peretz, I., Champod, S. & Hyde, K.(2003). Varieties of Musical Disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences , 999, 58-75*.
- Rothman, H. B., Rullman, J. F., & Arroyo, A. A. (1990). Inter-and intrasubject changes in vibrato: Perceptual and acoustic aspects. *Journal of Voice, 4, 309-316*.
- Russo, F. A., & Thompson, W. F. (2005). An interval size illusion: The influence of timbre on the perceived size of melodic intervals. *Perception & Psychophysics, 67*(4), 559-568.
- Sundberg, J. (2013). Perception of singing. In D. Deutsch (Ed.), *The psychology of music* (pp. 69–105). San Diego, CA: Academic Press.

References

- Sundberg, J., La, F. M., & Himonides, E. (in press). Intonation and Expressivity: A Single Case Study of Classical Western Singing. *Journal of voice*.
- Sundberg, J., Prame, E., & Iwarsson, J. (1996). Replicability and accuracy of pitch patterns in professional singers. In P. J. Davis & N. H. Fletcher (Eds.), *Vocal fold physiology, controlling complexity and chaos* (pp. 291–306). San Diego, CA: Singular Publishing Group.
- van Besouw, R. M., Brereton, J. S., & Howard, D. M. (2008). Range of tuning for tones with and without vibrato. *Music Perception, 26*, 145–155.
- Vurma, A., Raju, M., & Kuuda, A. (2010). Does timbre affect pitch?: Estimations by musicians and non-musicians. *Psychology of Music, 39*(3), 291-306.
- Vurma, A., & Ross, J. (2006). Production and perception of musical intervals. *Music Perception, 23*, 331–334.
- Warrier, C. M., & Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. Perception & Psychophysics, 64(2), 198-207.