Singing accuracy, listeners’ tolerance, and pitch analysis

Pauline Larrouy-Maestri
Pauline.Larrouy-Maestri@aesthetics.mpg.de

Johanna Devaney
Devaney.12@osu.edu
Musical errors

Contour error

Interval error

Tonality error
Musical errors

166 performances

http://sldr.org/sldr000774/en

Computer assisted method
(Larrouy-Maestri & Morsomme, 2013)
3 criteria

Judges

1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9
Out of tune

1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9
In tune

Larrouy-Maestri, P., Magis, D., Grabenhorst, M., & Morsomme, D. (revision). Layman or professional musician: Who makes the better judge?
Musical errors

• Intervals are important in the definition of vocal pitch accuracy in a melodic context

• When you are an experts, you pay attention to interval deviation and number of modulations

• But … tolerance?
Tolerance

- Pitch discrimination (e.g., [http://www.musicianbrain.com/pitchtest/](http://www.musicianbrain.com/pitchtest/))
- In a melodic context
  - Semitone (100 cents) Berkowska & Dalla Bella, 2009; Dalla Bella et al., 2007, 2009a, 2009b; Pfordresher & al., 2007, 2009, 2010
  - Quartertone (50 cents) Hutchins & Peretz; 2012; Hutchins, Roquet, & Peretz, 2012; Pfordresher & Mantell, 2014
- Tolerance of layman listeners for non-familiar melodies
  - Much less that a quartertone!
  - Whatever the type of error, the place and size of the interval
- Effect of expertise? Yes (most of the literature) No (Larrouy-Maestri et al., under revision)
# Tolerance: Participants

<table>
<thead>
<tr>
<th></th>
<th>Musicians</th>
<th>Non Musicians</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Gender</td>
<td>5 women</td>
<td>5 women</td>
</tr>
<tr>
<td>Age</td>
<td>( M = 41 ) (SD = 11.85)</td>
<td>( M = 41 ) (SD = 12)</td>
</tr>
<tr>
<td>Instrument</td>
<td>20 chords</td>
<td>no history of choral singing</td>
</tr>
<tr>
<td></td>
<td>11 wind</td>
<td>no formal musical training (max 2 years and no practice during the past 5 years)</td>
</tr>
<tr>
<td></td>
<td>4 percussions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 singers</td>
<td></td>
</tr>
<tr>
<td>Years of training</td>
<td>( M = 30.7 ) (SD = 12.32)</td>
<td></td>
</tr>
<tr>
<td>Starting</td>
<td>( M = 8.8 ) (SD = 4.63)</td>
<td>hearing threshold below 20 dB HL</td>
</tr>
<tr>
<td>Audiometry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Production task</td>
<td></td>
<td>ability to perform Happy Birthday with respect to appropriate melodic contour</td>
</tr>
<tr>
<td>MBEA (Peretz et al., 2003)</td>
<td></td>
<td>no deficit in music perception</td>
</tr>
</tbody>
</table>
Tolerance: Material

- **Familiar and Non-Familiar melodies**

- **Online questionnaire**
  - 399 participants from 13 to 70 years old ($M = 29.81$)
  - Familiarity ratings
    - $t(398) = 20.92, p < .001$
    - No effect of expertise on the ratings ($p > .05$)
Tolerance: Procedure

Methods of limits
(Van Besouw, Brereton, & Howard, 2008)

Two times Test-retest paradigm
Tolerance: Test-retest

- Highly significant correlation ($r(60) = .91, p < .001$)

- Training effect ($t(59) = 2.92, p = .005$)

- No effect of the direction of the deviation (i.e., enlargement vs. compression) $t(59) = -.96, p = .34$

- No effect of expertise ($p = .08$) or familiarity ($p = .71$) or interaction ($p = .65$) on the evolution test-retest
Tolerance: Effect of expertise and familiarity

- Effect of expertise \((F(1, 116) = 139.11, p < .001, \eta^2 = .54)\)
- No effect of familiarity \((F(1, 116) = 2.74, p = .10)\)
- No interaction \((F(1, 116) = .60, p = .44)\)
Tolerance: Effect of expertise and familiarity

- Low tolerance of all listeners when listening to melodies slightly out of tune (less than a quarter tone)

- Highly significant expertise effect, even for a familiar song well known by the participants (i.e., Happy Birthday)

- Training effect (mainly for the musicians)

- But … perceptual limit of musicians?
Pitch analysis
• Carl Seashore (1938) and colleagues studied timing, dynamics, intonation, and vibrato in pianists, violinists, and singers
  • Equipment: piano rolls, films of the movement of piano hammers during performance, phono-photographic apparati
Historical Methods
Phonophotography technique

- Frequency graphed in 10 cent units
- Intensity graphed in decibels
- Timing information as a function of linear space
Manual Annotation by Tapping

This webpage will auto correct reverse conducting taps so that they align with the nearest onset in the audio. The input tapping data can contain taps for all of the events, or just a selection of the events, such as the beats.

Input data is a text file with the event times in seconds on the first column of each line as output from Sonic Visualiser annotation layers. The tapping data is usually generated manually by tapping to a audio recording in Sonic Visualiser. The onset data is usually generated automatically from a plug-in for Sonic Visualiser, such as Spectral Reflux.

First, specify the location of the beat tapping data to be processed in one of the fields below.

Upload a file from your computer:

Or, paste the contents of the data file here:

Or, specify a data file URL:
Manual Annotation with Software

Audio Sculpt + Open Music

Visualisation of treatments in AudioSculpt software: segmentation and chord sequence analyses

Conversion in millicent

AS→OM

LISP length

Arithm-ser

Posn-match

Selection of the notes of the melody (œ, œ, œ…)

Musical representation of the analysis

Selection of the 1st partial of the chord

OMLOOP

List-minimum

file.txt

Results of the analysis
(6000 6200 6000 65000 6400 6000)

file.xls

Computation of errors
(pitch interval deviation, tonal centre deviation)
Manual Annotation with Software

PRAAT
Manual Annotation with Software

Audacity
Automatic Annotation

Sonic Visualiser
Automatic Annotation

TONY
Automatic Annotation

Melodyne
Identify Note Onsets and Offsets

Fundamental Frequency ($F_0$) Estimation

Perceived Pitch

Evolution of $F_0$
Score-guided performance data extraction

Monophonic and quasi-polyphonic

- Timing information is available via MIDI/audio alignment
- Fundamental frequency ($F_0$), and amplitude can be reliably extracted

Devaney, Mandel, and Ellis (2009)
Score-guided performance data extraction

Polyphonic

- Timing information (including asynchronies between lines) is available in the alignment
- \( F_0 \) and amplitude are harder to extract
- Currently exploring the using High Resolution methods with Roland Badeau for the task of score-guided extracting of frequency and loudness information in polyphonic audio

Devaney and Ellis (2009)
Devaney (2014)
## Perceived Pitch

### Possible calculation methods

<table>
<thead>
<tr>
<th>Study</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shonle and Horan (1980)</td>
<td>Geometric mean over the duration of the note</td>
</tr>
</tbody>
</table>
| Iwamiya, Kosugi, and Kitamura (1983)| • Center frequency between peaks and troughs in vibratos and symmetrical trills  
                                  | • In asymmetrical trills pitch shifts according to the direction of the asymmetry |
| D’Alessandro and Castellengo (1994, 1995) | • $F_0$ at the end of the note was more significant for the pitch perception than the beginning of the note. Mean of the steady-state portion of the note rather than the mid-point between the maximum and minimum frequencies |
| Gockel, Moore, and Carlyon (2001)   | • Weighted mean based on the fundamental frequencies’ rate of change, with higher weightings for frames that had a smaller rate of change |
Evolution of $F_0$

Modeling note trajectories

- Characterizing $F_0$ trajectories is under-studied
- One option is to decompose of $F_0$ trace with the Discrete Cosine Transform to estimate slope and curvature

Devaney, Mandel and Fujinaga (2011)
Devaney and Wessel (2013)
AMPACT
Automatic Music Performance and Comparison Toolkit

www.ampact.org
Thank you for your attention!

Johanna Devaney
Devaney.12@osu.edu

Pauline Larrouy-Maestri
Pauline.Larrouy-Maestri@aesthetics.mpg.de