

# Semitone Tuning in Vocal Performance

Johanna Devaney, Jason Hockman  
Jonathan Wild, Peter Schubert,  
Ichiro Fujinaga



McGill

CIR  
MMT

Centre for Interdisciplinary Research  
in Music Media and Technology

Fonds de recherche  
Société et culture  
Québec The logo for Québec's Fonds de recherche Société et culture, featuring the word "Québec" in a bold, black, sans-serif font next to a blue square containing two white fleur-de-lis symbols.



Social Sciences and Humanities  
Research Council of Canada

Conseil de recherches en  
sciences humaines du Canada

Canada The logo for the Government of Canada, featuring the word "Canada" in a bold, black, sans-serif font next to a small red Canadian flag icon.



# Introduction

Prior Work on Solo Vocal Intonation

Experiment on Two-Part Vocal Intonation

*Method - Results - Discussion*

Conclusions

# Introduction

## Prior Work on Solo Vocal Intonation

## Experiment on Two-Part Vocal Intonation

*Method - Results - Discussion*

## Conclusions

# Introduction

## Prior Work on Solo Vocal Intonation

## Experiment on Two-Part Vocal Intonation

*Method - Results - Discussion*

## Conclusions

# Introduction

Prior Work on Solo Vocal Intonation

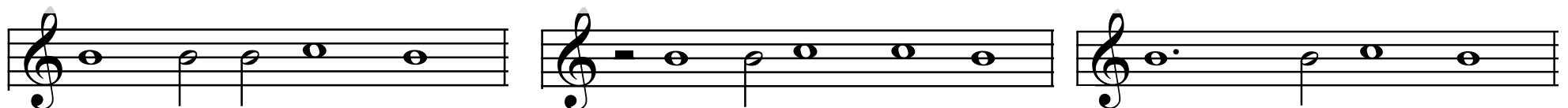
Experiment on Two-Part Vocal Intonation

*Method - Results - Discussion*

# Conclusions

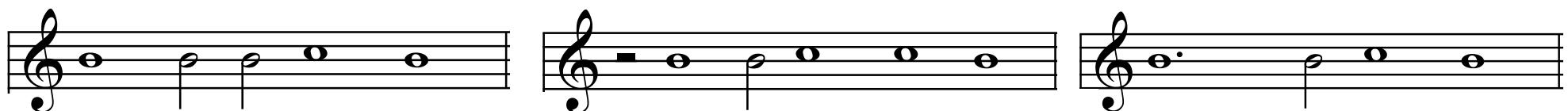
# Introduction

- Study considers the effect of accompaniment, de-tuning, and harmonic context on melodic semitone tuning
- Singers performed a single pattern (with three rhythms) with over a recorded accompaniment that consisted of 15 different patterns and 3 different tunings



# Introduction

- Study considers the effect of accompaniment, de-tuning, and harmonic context on melodic semitone tuning
- Singers performed a single pattern (with three rhythms) with over a recorded accompaniment that consisted of 15 different patterns and 3 different tunings



# Introduction

## Prior Work on Solo Vocal Intonation

## Experiment on Two-Part Vocal Intonation

*Method - Results - Discussion*

## Conclusions

# Prior Work on Intonation

- Schoen (1922) studied 5 accompanied professional singers
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Sundberg, Prame, and Iwarsson (1995) studied 10 accompanied professional singers
  - average difference between repeated notes was 8 cents
- Prame (1997) studied 10 accompanied professional singers
  - intonation deviated substantially, but not consistently, from equal temperament

# Prior Work on Intonation

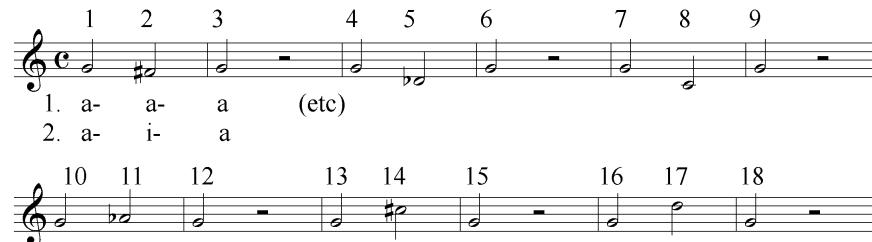
- Schoen (1922) studied 5 accompanied professional singers
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Sundberg, Prame, and Iwarsson (1995) studied 10 accompanied professional singers
  - average difference between repeated notes was 8 cents
- Prame (1997) studied 10 accompanied professional singers
  - intonation deviated substantially, but not consistently, from equal temperament

# Prior Work on Intonation

- Schoen (1922) studied 5 accompanied professional singers
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Sundberg, Prame, and Iwarsson (1995) studied 10 accompanied professional singers
  - average difference between repeated notes was 8 cents
- Prame (1997) studied 10 accompanied professional singers
  - intonation deviated substantially, but not consistently, from equal temperament

# Prior Work on Intonation

- Vurma and Ross (2006) studied 13 *a cappella* professional singers
  - ascending/descending semitones smaller than EQT
  - ascending/descending fifths larger than than EQT



- Vurma (2010) studied 15 accompanied professional singers
  - singers' intonation did not change significantly when the synthesized voice was detuned 20 or 40 cents narrower



# Prior Work on Intonation

- Vurma and Ross (2006) studied 13 *a cappella* professional singers
  - ascending/descending semitones smaller than EQT
  - ascending/descending fifths larger than than EQT



- Vurma (2010) studied 15 accompanied professional singers
  - singers' intonation did not change significantly when the synthesized voice was detuned 20 or 40 cents narrower



# Prior Work on Intonation

- Devaney, Mandel, and Fujinaga (2011)
  - Schubert's "Ave Maria"
    - 6 undergraduate singers
    - 6 professional singers
  - Impact of training on tuning of melodic semitones
    - pros were closer to EQT, non-pros were flatter
    - non-pros exhibited a difference between accompanied and *a cappella* performances
    - ascending semitones were 7–8 cents larger on avg.
    - non-pros tended to compress leading tones
    - pros' semitones were 6 cents larger on average
    - pros were more consistent with one another

# Introduction

## Prior Work on Solo Vocal Intonation

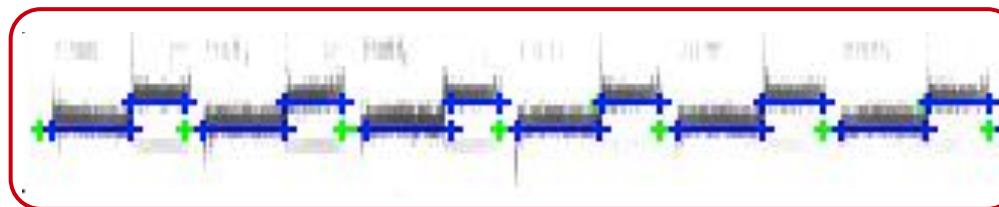
## Experiment on Two-Part Vocal Intonation

*Method - Results - Discussion*

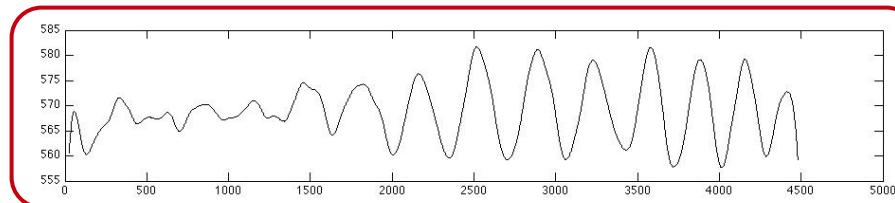
## Conclusions

# Method: *Extracting Data*

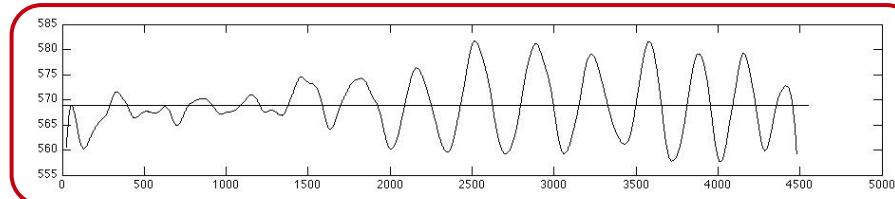
Identify Note Onsets and Offsets



Fundamental Frequency (Fo) Estimation



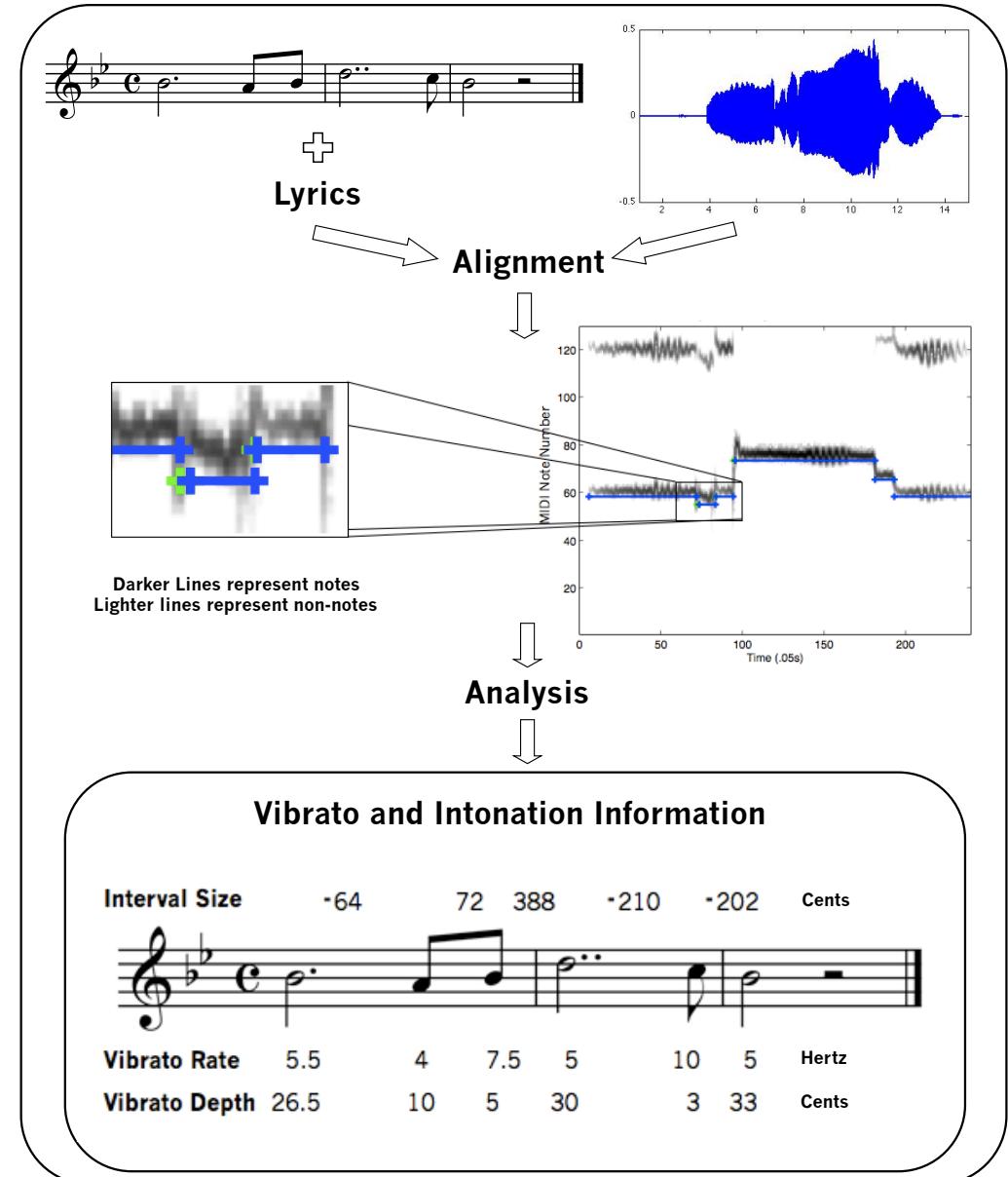
Perceived Pitch



# Method: AMPACT

Automatic Music  
Performance Analysis  
and Comparison Toolkit

[www.ampact.org](http://www.ampact.org)



# Method: *Participants*

- 6 singers
  - 3 professional singers
    - average age: 42 ( $SD=10$ )
    - years of private voice lessons: 9 ( $SD=2$ )
    - years of daily practice: 14 ( $SD=6$ )
    - hours of weekly rehearsal time: 10 ( $SD=4$ )
  - 3 non-professional singers
    - average age: 53 ( $SD=0.5$ )
    - years of private voice lessons: 0.67 ( $SD=0.5$ )
    - years of daily practice: 7 ( $SD=6$ )
    - hours of weekly rehearsal time: 3 ( $SD=2.5$ )

# Method: *Detuning Scheme*

Just Intonation					
1	D	G	D		
	-2	-4	-2		
2	D	D	F		
	-2	-2	14		
3	D	Bb	F		
	-2	12	14		
4	F	D	A		
	14	-2	0		
5	F	G	A		
	14	18	0		
6	F	G	D		
	14	18	20		
7	F	Eb	F		
	-8	-12	-8		
8	A	G	F		
	0	-4	-8		
9	A	D	F		
	0	-2	-8		
10	F	Eb	D	C	F
	14	10	-2	16	14
11	F	G	C	D	
	-8	-4	-6	-2	
12	F	Bb	F		
	-8	-10	-8		
13	D	G	A	A	
	20	18	22	22	
14	G	F	E	F	
	-4	-8	2	-8	
15	G	F	Eb	F	
	-4	-8	-12	-8	

Modified Just Intonation					
D	G	D			
-2	18	20			
D	D	F			
-2	-25	-8			
D	Bb	F			
-2	-10	-8			
F	D	A			
14	20	22			
F	G	A			
-8	-4	-23			
F	G	D			
-8	-27	-25			
F	Eb	F			
14	33	14			
A	G	F			
0	18	37			
A	D	F			
22	20	14			
F	Eb	D	C	F	
-8	-12	-2	-6	-8	
F	G	C	D		
14	18	16	20		
F	Bb	F			
14	12	14			
D	G	A	A		
	-27	-23	-23		
G	F	E	F		
18	14	2	14		
G	F	Eb	F		
18	14	10	14		

# Method: *Exercises*

A musical staff in G clef and common time. It consists of four measures. Red circles highlight specific intervals: a minor third (m3) in the first measure, a major sixth (m6) in the second measure, a perfect fifth (P5) in the third measure, and another major sixth (m6) in the fourth measure.

A musical staff in G clef and common time. It consists of four measures. Red circles highlight specific intervals: two minor thirds (m3) in the first and second measures, a perfect fifth (P5) in the third measure, and another minor third (m3) in the fourth measure.

A musical staff in G clef and common time. It consists of four measures. Red circles highlight specific intervals: a major sixth (m6) in the first measure, a perfect fifth (P5) in the second measure, a minor third (m3) in the third measure, and a perfect eighth (P8) in the fourth measure.

A musical staff in G clef and common time. It consists of three measures. Red circles highlight specific intervals: a minor third (m3) in the first measure, a tritone (TT) in the second measure, and a perfect fifth (P5) in the third measure.

# Results

- Overall average size of semitones: 85 cents ( $SD=21$ )
- Ascending semitones average size: 94 cents ( $SD=17$ )
- Descending semitones average size: 76 cents ( $SD=24$ )
- Factors considered in analysis of variance (ANOVA)
  - Direction
  - Pro/Non-Pro
  - Pitch Height (starting on G or B)
  - Tuning (EQT, Just, Modified Just)
  - Vertical Interval (m3, TT, P5, m6, P8)

# Results

- DIRECTION: Ascending semitones were on average **21 cents** larger on average than descending semitones
- GROUP DIFFERENCES: Non-pros' semitones were on average 17 cents on average smaller than pros' semitones
- PITCH HEIGHT: Semitones in exercises starting on G were 6 cents larger on average than the semitones than exercises starting on Bb
- TUNING: *not significant*
- VERTICAL INTERVALS: Semitones sung a perfect octave above the lower voice were 7 cents larger on average than those sung above other intervals
  - *there were no significant differences for other intervals*

# Results

- **DIRECTION:** Ascending semitones were on average 21 cents larger on average than descending semitones
- **GROUP DIFFERENCES:** Non-pros' semitones were on average 17 cents on average smaller than pros' semitones
- **PITCH HEIGHT:** Semitones in exercises starting on G were 6 cents larger on average than the semitones than exercises starting on Bb
- **TUNING:** *not significant*
- **VERTICAL INTERVALS:** Semitones sung a perfect octave above the lower voice were 7 cents larger on average than those sung above other intervals
  - *there were no significant differences for other intervals*

# Results

- DIRECTION: Ascending semitones were on average 21 cents larger on average than descending semitones
- GROUP DIFFERENCES: Non-pros' semitones were on average 17 cents on average smaller than pros' semitones
- PITCH HEIGHT: Semitones in exercises starting on G were 6 cents larger on average than the semitones than exercises starting on Bb
- TUNING: *not significant*
- VERTICAL INTERVALS: Semitones sung a perfect octave above the lower voice were 7 cents larger on average than those sung above other intervals
  - *there were no significant differences for other intervals*

# Results

- DIRECTION: Ascending semitones were on average 21 cents larger on average than descending semitones
- GROUP DIFFERENCES: Non-pros' semitones were on average 17 cents on average smaller than pros' semitones
- PITCH HEIGHT: Semitones in exercises starting on G were 6 cents larger on average than the semitones than exercises starting on Bb
- TUNING: *not significant*
- VERTICAL INTERVALS: Semitones sung a perfect octave above the lower voice were 7 cents larger on average than those sung above other intervals
  - *there were no significant differences for other intervals*

# Results

- DIRECTION: Ascending semitones were on average 21 cents larger on average than descending semitones
- GROUP DIFFERENCES: Non-pros' semitones were on average 17 cents on average smaller than pros' semitones
- PITCH HEIGHT: Semitones in exercises starting on G were 6 cents larger on average than the semitones than exercises starting on Bb
- TUNING: *not significant*
- VERTICAL INTERVALS: Semitones sung a perfect octave above the lower voice were 7 cents larger on average than those sung above other intervals
  - *there were no significant differences for other intervals*

# Discussion

- Schoen (1922)
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Vurma and Ross (2006)
  - ascending/descending semitones smaller than EQT
- Vurma (2010)
  - singers' intonation did not change significantly when the synthesized voice was detuned
- Devaney, Mandel, and Fujinaga (2011)
  - ascending semitones larger on average
  - non-pros' semitones smaller on average

# Discussion

- Schoen (1922)
  - sharper than equal temperament 
  - less sharp when descending than when ascending
- Vurma and Ross (2006)
  - ascending/descending semitones smaller than EQT
- Vurma (2010)
  - singers' intonation did not change significantly when the synthesized voice was detuned
- Devaney, Mandel, and Fujinaga (2011)
  - ascending semitones larger on average
  - non-pros' semitones smaller on average

# Discussion

- Schoen (1922)
  - sharper than equal temperament 
  - less sharp when descending than when ascending 
- Vurma and Ross (2006)
  - ascending/descending semitones smaller than EQT
- Vurma (2010)
  - singers' intonation did not change significantly when the synthesized voice was detuned
- Devaney, Mandel, and Fujinaga (2011)
  - ascending semitones larger on average
  - non-pros' semitones smaller on average

# Discussion

- Schoen (1922)
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Vurma and Ross (2006)
  - ascending/descending semitones smaller than EQT
- Vurma (2010)
  - singers' intonation did not change significantly when the synthesized voice was detuned
- Devaney, Mandel, and Fujinaga (2011)
  - ascending semitones larger on average
  - non-pros' semitones smaller on average

# Discussion

- Schoen (1922)
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Vurma and Ross (2006)
  - ascending/descending semitones smaller than EQT
- Vurma (2010)
  - singers' intonation did not change significantly when the synthesized voice was detuned
- Devaney, Mandel, and Fujinaga (2011)
  - ascending semitones larger on average
  - non-pros' semitones smaller on average



# Discussion

- Schoen (1922)
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Vurma and Ross (2006)
  - ascending/descending semitones smaller than EQT
- Vurma (2010)
  - singers' intonation did not change significantly when the synthesized voice was detuned
- Devaney, Mandel, and Fujinaga (2011)
  - ascending semitones larger on average
  - non-pros' semitones smaller on average

# Discussion

- Schoen (1922)
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Vurma and Ross (2006)
  - ascending/descending semitones smaller than EQT
- Vurma (2010)
  - singers' intonation did not change significantly when the synthesized voice was detuned
- Devaney, Mandel, and Fujinaga (2011)
  - ascending semitones larger on average
  - non-pros' semitones smaller on average



# Discussion

- Schoen (1922)
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Vurma and Ross (2006)
  - ascending/descending semitones smaller than EQT
- Vurma (2010)
  - singers' intonation did not change significantly when the synthesized voice was detuned
- Devaney, Mandel, and Fujinaga (2011)
  - ascending semitones larger on average
  - non-pros' semitones smaller on average

# Discussion

- Schoen (1922)
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Vurma and Ross (2006)
  - ascending/descending semitones smaller than EQT
- Vurma (2010)
  - singers' intonation did not change significantly when the synthesized voice was detuned
- Devaney, Mandel, and Fujinaga (2011)
  - ascending semitones larger on average ✓
  - non-pros' semitones smaller on average

# Discussion

- Schoen (1922)
  - sharper than equal temperament
  - less sharp when descending than when ascending
- Vurma and Ross (2006)
  - ascending/descending semitones smaller than EQT
- Vurma (2010)
  - singers' intonation did not change significantly when the synthesized voice was detuned
- Devaney, Mandel, and Fujinaga (2011)
  - ascending semitones larger on average
  - non-pros' semitones smaller on average

# Introduction

Prior Work on Solo Vocal Intonation

Experiment on Two-Part Vocal Intonation

*Method - Results - Discussion*

# Conclusions

# Conclusions

- This study verified findings in earlier studies
  - Bigger difference in the experience levels of pros and non-pros
- This study contributed new findings for influence of pitch height and vertical interval on melodic semitone intonation
- Future Work
  - Explore the effect of pitch height in more detail
  - Examine whole tone intonation
  - Use exercises with sequential ascending and descending intervals

# Acknowledgements



McGill

CIR  
MMT

Centre for Interdisciplinary Research  
in Music Media and Technology

*Fonds de recherche*  
*Société et culture*

Québec



Social Sciences and Humanities  
Research Council of Canada

Conseil de recherches en  
sciences humaines du Canada

Canada

# Acknowledgements



McGill

CIR  
MMT

Centre for Interdisciplinary Research  
in Music Media and Technology



Social Sciences and Humanities  
Research Council of Canada

Conseil de recherches en  
sciences humaines du Canada

Canada

# Thank you!

# References

- de Cheveigné A. and Kawahara H. 2002. YIN: A fundamental frequency estimator for speech and music. *Journal of the Acoustical Society of America*, 111, 1917–30.
- Devaney, J., M. Mandel, I., D. P. W. Ellis, and I. Fujinaga. 2011. Automatically extracting performance data from recordings of trained singers. *Psychomusicology: Music, Mind and Brain* 21 (1–2): 108–36.
- Gockel H., Moore B., and Carlyon R. 2001. Influence of rate of change of frequency on the overall pitch of frequency-modulated tones. *Journal of the Acoustical Society of America*, 109, 701–12.
- Prame E. 1997. Vibrato extent and intonation in professional western lyric singing. *Journal of the Acoustical Society of America*, 102, 616–21.
- Schoen M. 1922. An experimental study of the pitch factor in artistic singing. *Psychological Monographs*, 31, 230–59.
- Sundberg, J., E. Prame, and J. Iwarsson. 1995. Replicability and accuracy of pitch patterns in professional singers. *STL-Quarterly Progress and Status Report* 36 (2–3): 51–62.

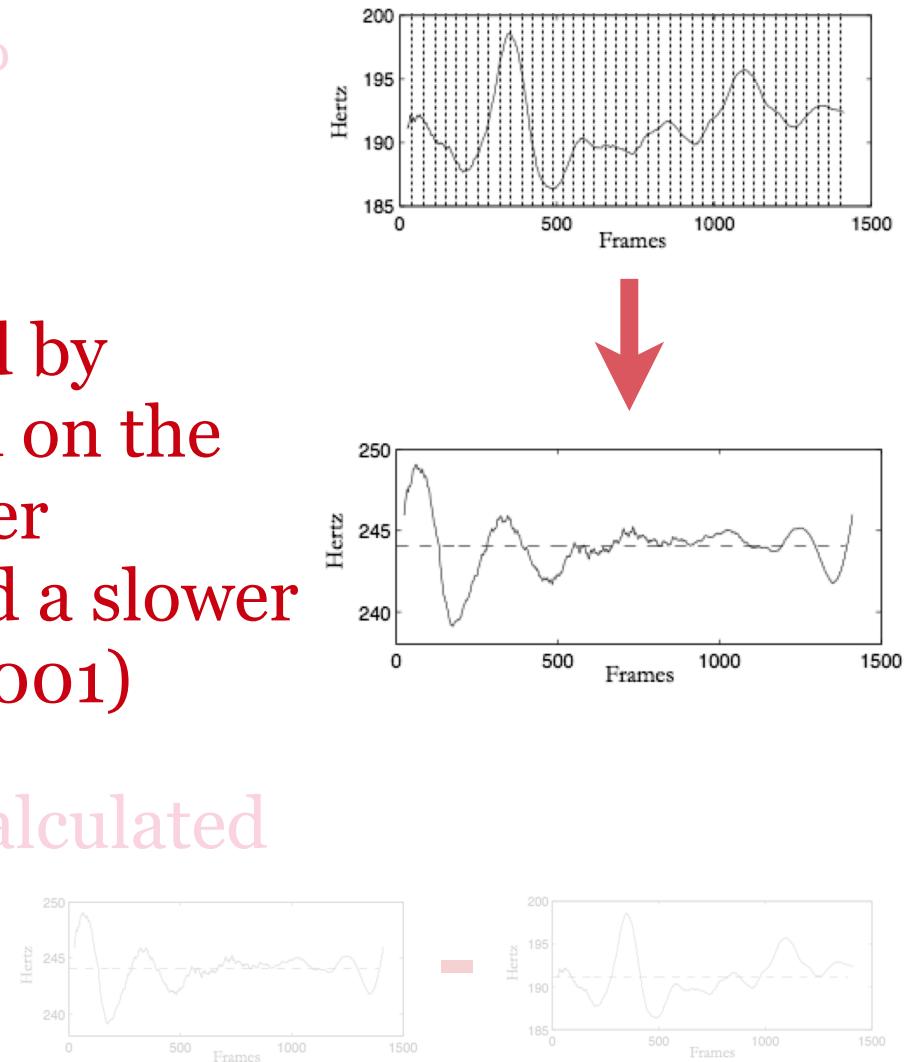
# Method: $F_o$ and Interval

- The YIN algorithm used for  $F_o$  estimation (de Cheveigné and Kawahara 2002)
- Perceived pitch was calculated by taking a weighted mean based on the  $F_o$ 's rate of change, with higher weightings for frames that had a slower rate of change (Gockel et al. 2001)
- Horizontal interval size was calculated as the difference between two perceived pitch calculations



# Method: $F_o$ and Interval

- The YIN algorithm used for  $F_o$  estimation (de Cheveigné and Kawahara 2002)
- Perceived pitch was calculated by taking a weighted mean based on the  $F_o$ 's rate of change, with higher weightings for frames that had a slower rate of change (Gockel et al. 2001)
- Horizontal interval size was calculated as the difference between two perceived pitch calculations



# Method: $F_o$ and Interval

- The YIN algorithm used for  $F_o$  estimation (de Cheveigné and Kawahara 2002)
- Perceived pitch was calculated by taking a weighted mean based on the  $F_o$ 's rate of change, with higher weightings for frames that had a slower rate of change (Gockel et al. 2001)
- Horizontal interval size was calculated as the difference between two perceived pitch calculations

