The influence of sung vowels on pitch perception

Johanna Devaney and Derek Richardson
School of Music, The Ohio State University

Introduction
Background and Motivations

1

Experiment
Methods, Results, and Discussion

2

Conclusions
Future Directions and Summary

3

Introduction

Fundamental Research Question

- **The relationship between pitch and vowels has been studied extensively for speech, and to a lesser extent, singing**
  - Research on speech has address both production and perception
  - Research on singing has focused on production
- **Do different sung vowels with the same mean F_0 elicit different pitch percepts in listeners?**

Background

Production Experiments on Vowels

- **in speech there is an intrinsic pitch of vowels (IPV)**
  - different spoken vowels are consistently produced at different pitch heights (Sapir 1989)
- **The same effect has been observed in singing, but to a lesser degree**
  - raised F_0 in front vowels /i/ and /y/ (Ternstrom, Sundberg, and Colldén 1988) – *choral singers*
  - higher F_0 for /i/ than for /a/ (Fowler and Brown, 1997) - *non-singers*

Background

Perception Experiments on Vowels

- **Fowler and Brown (1997) also ran a non-adaptive forced-choice perception experiment with synthesized sung tones**
  - A Front/Close vowel (/i/) was perceived as higher than a Back/Open vowel (/a/) when the mean F_0 was the same

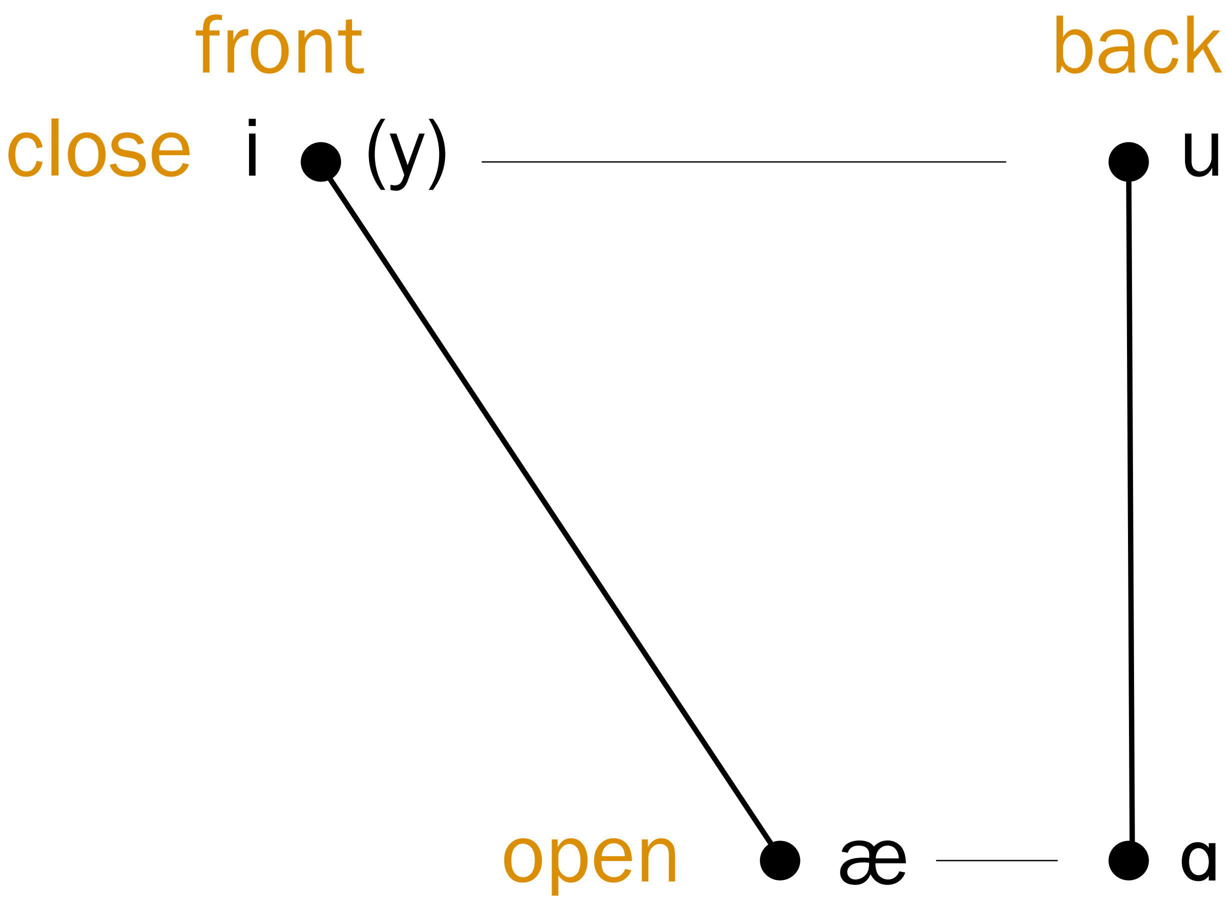
Motivation

Developing a vowel-contextualized pitch perception model

- **Robust models of perceived pitch are necessary to accurately study tuning in singing**
- **Extensive work has been done on the pitch of stable vibrato tones**
  - (d’Alessandro and Castellengo, 1994; Gockel, Moore, and Carlyon, 2001)
- **And some work has been done on glides**
  - (d'Alessandro, Rosset, and Piot, 1995)
- **But there are no models for different sung vowels**
  - due to a lack of comprehensive data

Method

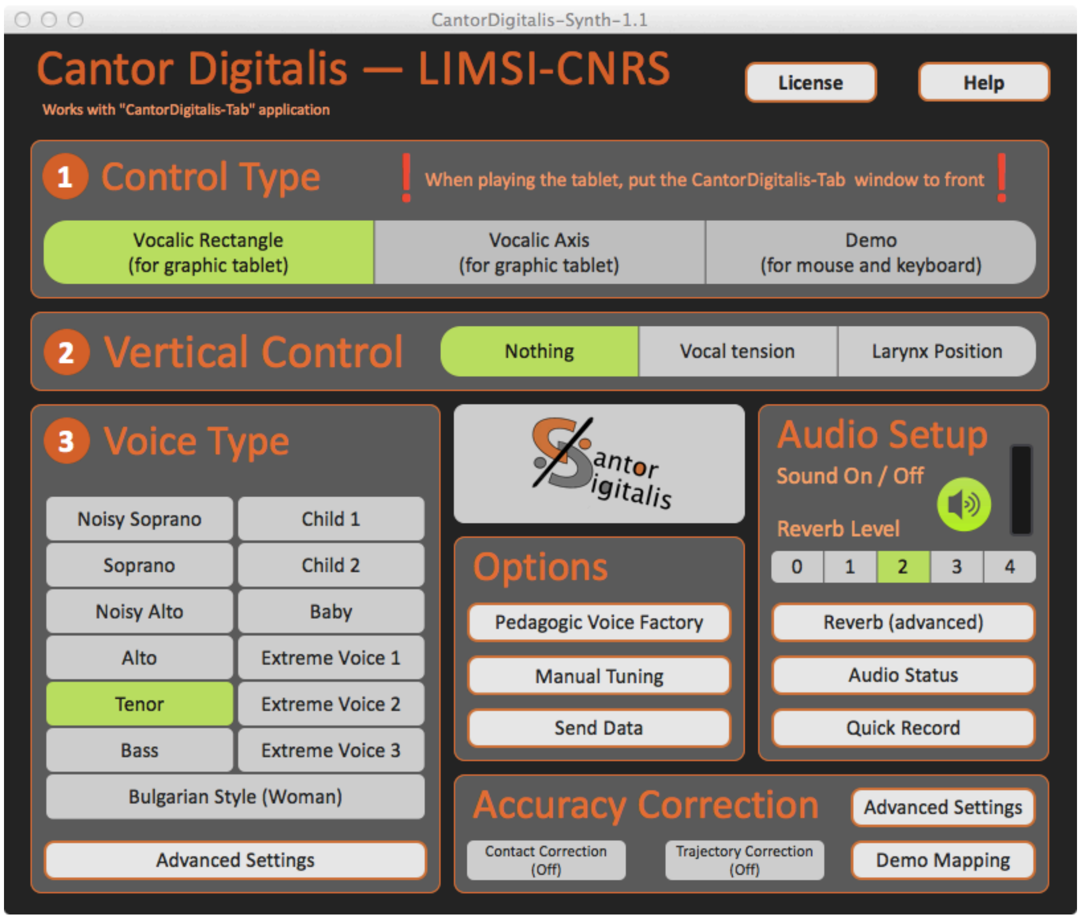
Vowels under consideration



	F1	F2
i	290	1900
u	287	750
æ	690	2210
ɑ	650	1200

Method

Stimulus Design



http://cantordigitalis.limsi.fr

(d'Alessandro et al. 2014)

Isolated vowels

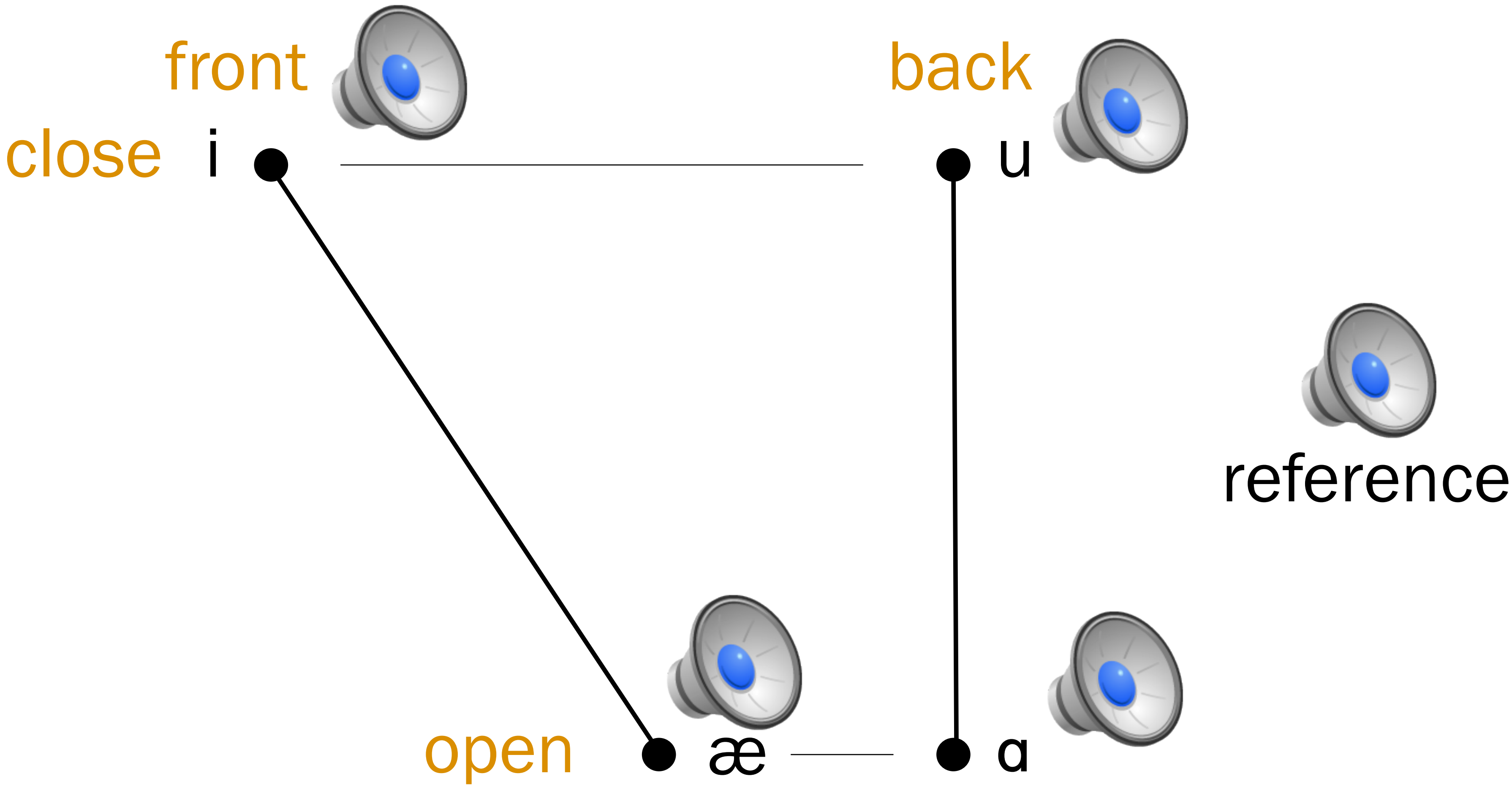
500 ms
- 100 ms onset ramp
- 300 ms steady tone
- 100 ms offset ramp

130.8 Hz (C_3)
- +/- 100 cents
- F0 verified with YIN (de Cheveigné and Kawahara 2002)

Reference stimuli
- same synthesis
- no formants

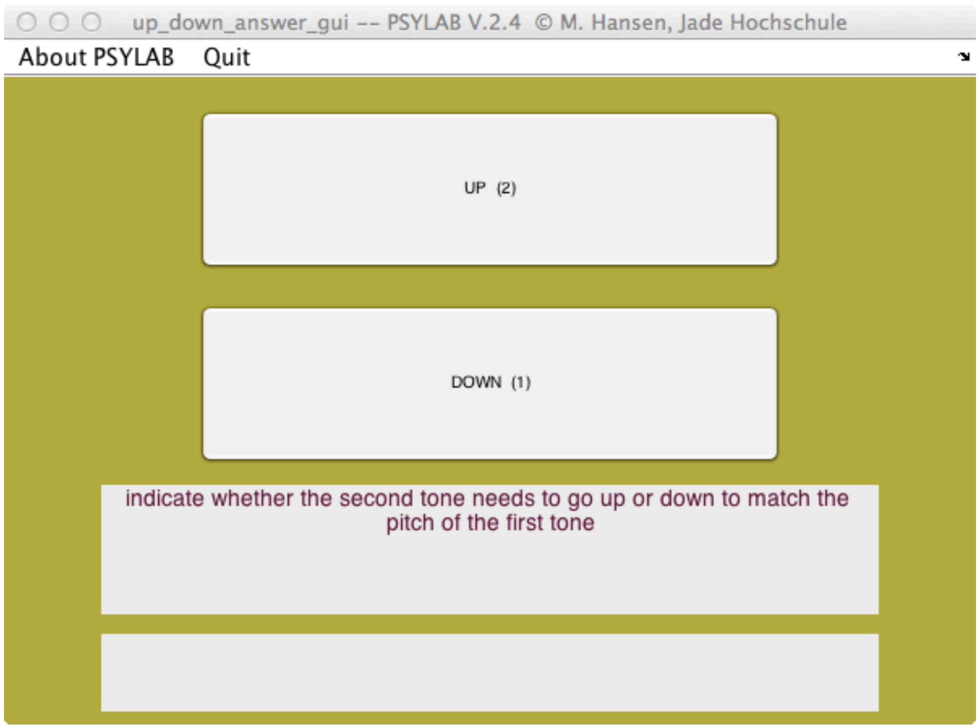
Method

Stimulus Design



Method

Experimental Design



http://www.hoertechnik-audiologie.de/psylab/

Matching paradigm

- Start: +33 cents
  - Halved subsequently
- Reference tone heard first
- Stop criterion: 2 reversals

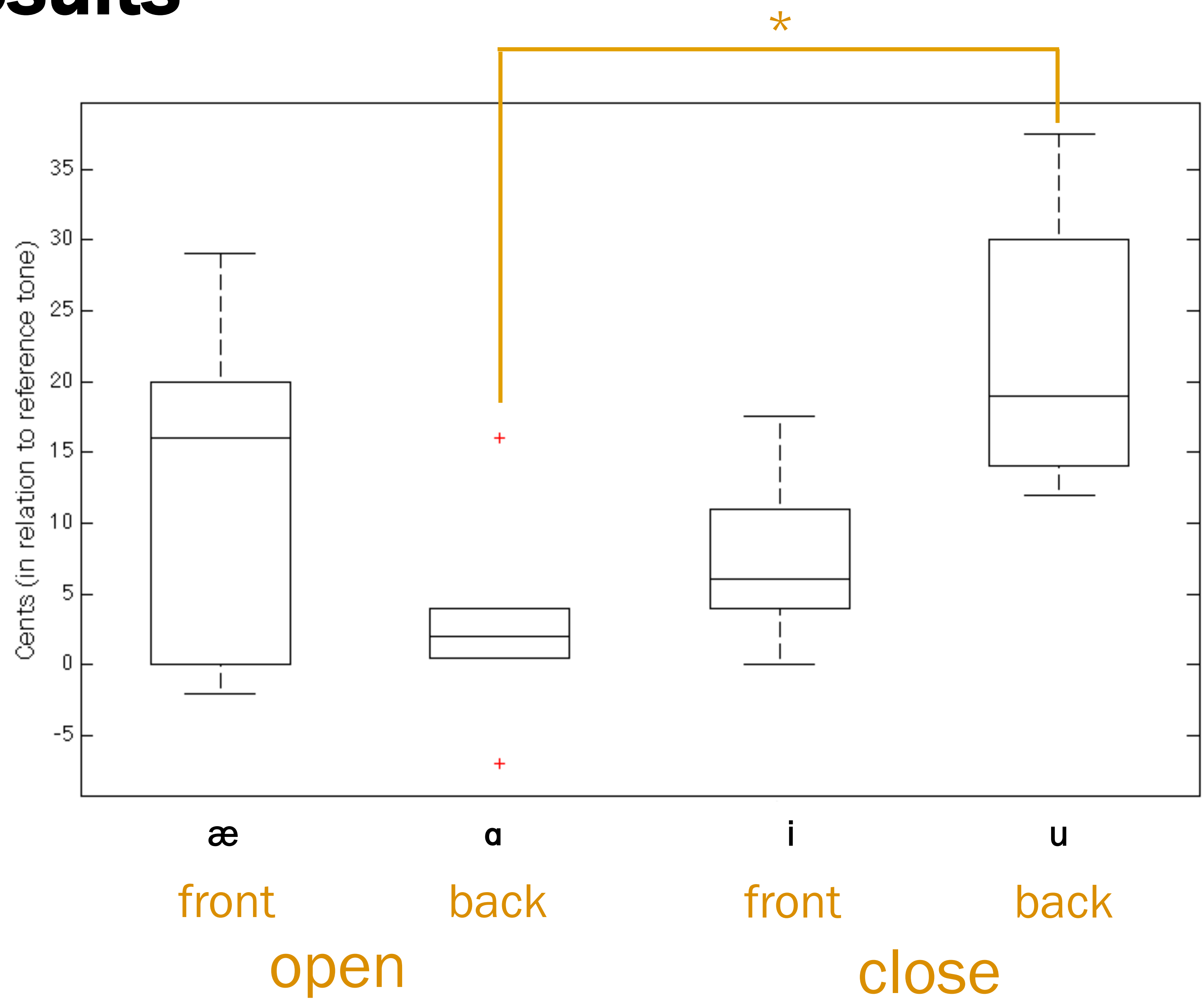
12 trials (3 each)

- First 4 for acclimatization

4 subjects (3 female, 1 male)

- 3 with PhDs in audio-related field and 1 PhD student

Results



*
Cents (in relation to reference tone)
35
30
25
20
15
10
5
0
-5
æ
ɑ
i
u
front
back
front
back
open
close

Discussion

- **Back/Close (/u/) versus Back/Open (/ɑ/)**
  - -19 cents ($p > 0.001$)
- **Front/Close (/i/) versus Back/Open (/ɑ/)**
  - Fowler and Brown (1997): a Front/Close vowel (/i/) was perceived as higher than a Back/Open vowel (/a/) when the mean F_0 was the same
  - +14.5 (not sig, $p = 0.06$)
- **The differences in amounts of variance may be of interest but these data do not allow for a robust investigation of this**

Introduction

Background and Motivations

1

Experiment

Methods, Results, and Discussion

2

Conclusions

Future Directions and Summary

3

Conclusions

Future Directions

- **Run more subjects**
  - Musicians and non-musicians
  - Different native languages (Pape, 2005)
- **Expand to more vowels**
- **Evaluate the effect of duration**
- **Implement relevant findings into a perception model**

Conclusions

Summary

- This talk has presented an experiment on the perceived pitch of four extremal vowels
- The Back/Open vowel (/ɑ/) was perceived as significantly lower than the Back/Close vowel (/u/)
- We replicate the findings in Fowler and Brown (1997) regarding Front/Close vowels being perceived as higher than Back/Open vowels

Thank you!

References

d’Alessandro, C., & Castellengo, M. (1994). The pitch of short-duration vibrato tones. *JASA*, 95(3), 1617–30.

d’Alessandro, C. Feugère, L., Le Beux, S., Perrotin, O., & Rilliard, A. (2014) Drawing melodies: Evaluation of Chironomic Singing Synthesis. *JASA*, 135(6), 3601–12.

d'Alessandro, C., Rosset, S., & Piot, O. (1995). Measurement of pitch perception for F0 glides In the *Proceedings of EuroSpeech*, 957–60.

de Cheveigné, Alain, & Kawahara, Hideki. (2002). YIN, a fundamental frequency estimator for speech and music. *JASA*, 111(4), 1917–1930.

Fowler, C., & Brown, J. (1997). Intrinsic F0 differences in spoken and sung vowels and their perception by listeners. *Perception & Psychophysics, 59*(5), 729–38.

Gockel, H., Moore, B., & Carlyon, R. (2001). Influence of rate of change of frequency on the overall pitch of frequency-modulated tones. *JASA, 109*(2), 701–12.

Pape, D. (2005). Is pitch perception and discrimination of vowels language-dependent and influenced by the vowels spectral properties. In the *Proceedings of the International Conference on Auditory Display*, 340–3.

Sapir, S. (1989). The intrinsic pitch of vowels: theoretical, physiological, and clinical considerations. *Journal of Voice, 3*(1), 44–51.

Ternström, S., Sundberg, J., & Colldén, A.. (1988). Articulatory F0 Perturbations and Auditory Feedback. *Journal of Speech and Hearing Research, 31* 187–192.