Intonation Accuracy and Pitch Stability During Crescendo as the Voice Quality and Singer's Experience Indicator Among Choral Singers

Mateusz GAWLIK
AGH University of Science and Technology,
30–059 Cracow, Mickiewicza 30,
matgawel@agh.edu.p1

Wiesław WSZOŁEK
AGH University of Science and Technology,
30–059 Cracow, Mickiewicza 30,
wwszolek@agh.edu.pl

Abstract

The ability to correctly reproduce notes by the voice is one of the essential features of the singing task and called intonation. In combination with other parameters like timbre, formants, and sound attack, it affects the reception of listening impressions. In this paper, we present results of the examination concerning the automatic evaluation of intonation among the nonsingers, untrained and trained choral singers. We performed both pitch error during vocalization and pitch stability in crescendo task analysis among studied groups. We used Zero Band Filtering method to determine fundamental frequency from the singing signal. We noticed significant differences between singers with different skills and experience, and the possibility to classify the level of advancement of the singer by using intonation characteristic.

Keywords: intonation, singing voice, pitch accuracy, jitter, signal processing

1. Introduction

Nonprofessional singers tend to get common disorders due to too high laryngeal muscles tension or other reasons connected with unskillful singing. Lack of singing technique leads to vocal disorders that can cause the loss of the ability to sing or even a complete loss of voice. Correct determination and elimination of incorrect techniques is therefore essential task during singing training. The researchers of the singing voice assessment analyzed it in various ways. It is possible to assess advancement level by sound pressure level and phonation threshold pressure [1-3], subglottal pressure [4], glottal flow parameterization [5, 6], formants, and timbre evaluation [7, 8]. One of the factors useful for assessing the technical skills of a singer is also intonation [9]. Music experts can judge the sound as purely based only on pitch accuracy. Intonation is also an important factor when determining the singer's tessitura.

In [10] the researchers used intonation and vibrato parameter to classify singing technique into two groups (poor or good) with accuracy to the semitone. Authors used low–pass filtering to remove f0 fluctuations (connected with vibrato). They observed that pitch interval accuracy was constant and close to zero for good singers and irregular, with a positive trend to 100 cents for poor ones.

In [11] authors defined the intonation as the ability to reflect the pattern tone accurately. The singers were classified into four classes: unsatisfactory, satisfactory, good and very good, based on the correctly performed number of phrases according to pitch and slope values. They set up an acceptable range of tone within a deviation of 20 cents and the slope absolute value < 0.01. They noticed that poor singers tend to overstate or understate the pitch in certain sound intervals.

Researchers of the singing voice often use the standard deviation of the f0 as an indicator of pitch accuracy [12]. In [13] the average of the standard deviation was no greater than 15 cents for the different tones among good singers of an amateur choir.
It took about 150 ms for a singer to adjust to a heard change [14], where the ability of skilled choir members to adapt to sudden tones shifting by 50 or 100 cents. Those results are similar to [15], where rapid pitch changes of 120 ms duration at the beginning and end of a tone were observed.

The study [16] analysed pitch accuracy among professional singers during ascending and descending intervals. Their results correspond to [12] and proved that, on average, the singer's intonation was comparable to the patterns with the standard deviation of 20 cents.

Table 1.Singers' experience level and middle tones for different voice types. NS – nonsingers, US – untrained singers, TS – trained singers

NS	US	TS	Pattern tones	
p3	p2	p1	Bass	E2
p4	p6	p5	Baritone	G3
p7	p8	p13	Tenor	B3
p15	p9	p17	Alto	E4
p21	p10	p20	Mezzo	G4
	p11		Soprano	B4
	p12			
	p14			
	p16			
	p17			
	p18			
	p19			

There are a lot of studies related to intonation evaluation, here we reviewed only some of them due to article volume limitations. A detailed description of the current state of the art is included in [12] and [15].

Previous papers described intonation analyzes among singers at a similar level of skills or years of experience. We noticed that authors focus on a wider research group rarely. This observation encouraged us to answer the question: Is it possible to distinguish singers' advancement or experience level among nonsingers, untrained singers and trained singers by f0 analyzing, based on pitch accuracy? Additionally, based on f0 and vocal intensity relationships observed in [17] we hypothesized that pitch stability during crescendo is connected with experience and vocal training.

2. Material and Methods

2.1. *Data acquisition*

Sixteen choir members with different singing skills, two students of theatre academy and three ordinary students were recorded in the anechoic chamber at the University of Science and Technology in Cracow. We used G.R.A.S 40 AF high–precision condenser microphone, always positioned at the 50 cm distance from the singer's mouth. The receiver was connected to G.R.A.S 12AA 2–channel power module and M–Audio PROFIRE 610 audio interface with Octanes preamplifiers. We posted information about the studied group in Table 1.

2.2. *Recording session*

We asked singers to perform crescendo task at vowel *a* at the certain pitch. Table 1 contains detailed information about tones, used for that exercise. We divided the studied group into three subgroups, according to the level of experience:

- Nonsingers (NS) – people without any singing experience
- Untrained singers(US) – members of Cracow choirs with different singing skills and experience
- Trained singers (TS) – graduates of the vocal music academy.

2.3. *Signal parametrization*

We looked for useful parameters to find differences in intonation in nonsingers and singers, as well as if the differences occur for singers group on various advancement levels. First, we calculated fundamental frequency (f0) of the signal, as the primary intonation indicator. We also appointed additional parameters such as jitter and linear regression coefficient. We used frame-based analysis of the signal with 25 ms frames length and 50% overlapping. Fundamental frequency was obtained by Zero Band Filtering method [18]. We used the Kruskal–Wallis test to check whether differences between groups are statistically significant. It was used in because of non–normality of the distribution of the parameters.

3. Results

At first, we calculated pitch error among all examined groups (Fig. 1 left) as the absolute value of difference from the pattern pitch. The results are presented in Tab. 2. The results are not surprising, and most of the tested singers (except p3, p9, and p16) intoned pattern pitch with high accuracy (low pitch error). However, we found a statistically significant difference between the median values of each group for all vowels (Kruskal–Wallis test, $p < 0.01$). Additionally, trained singers obtained the lowest values of medians of pitch error for almost all vowels that ensures frequency resolution range 10–25 cents, which is similar to [11].

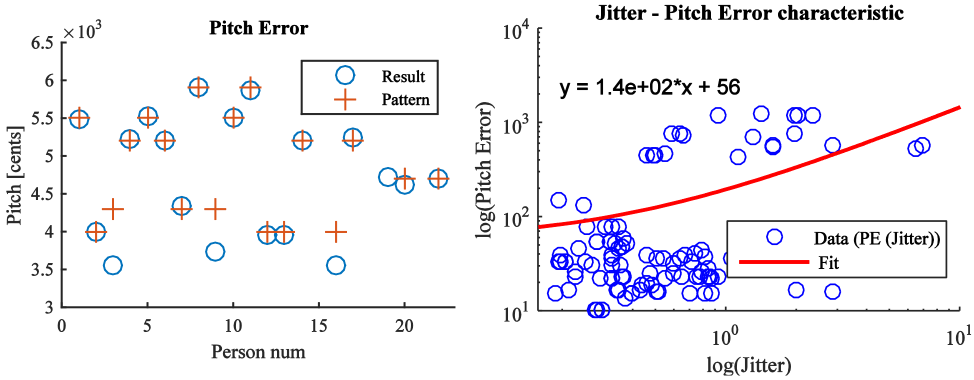

Figure 1. Pitch errors between patterns and sang vowels (left),
boxplots median of linear regression coefficients from *a* vowel during (right)

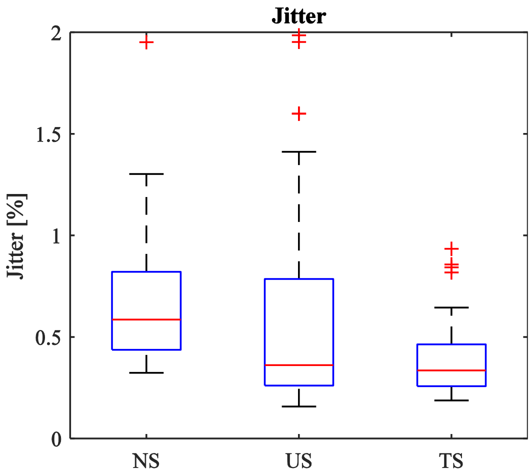

Figure 2. Boxplots median of linear regression coefficients from a vowel during crescendo task. NS – nonsingers, US – untrained singers, TS – trained singers

Next, the pitch error was expressed as a function of jitter parameter (Fig. 1 right). We observed that an increase of pitch error is connected with higher jitter values. Therefore untrained singers obtained higher values of this parameter than other groups (Fig. 3), but there was no statistically significant difference between these groups (Kruskal–Wallis test, $p > 0.1$). Finally, we carried out a trend line examination during crescendo sounds. In all studied cases, untrained and trained singers obtained lower values of linear regression coefficients than nonsingers (Figs. 4 – 7). The median values were presented in Tab. 3. Additionally, we observed oscillation in f0 characteristics in singers groups, which was not present in nonsingers (Fig. 6). Moreover, the oscillating trend is more regularly for trained, then untrained singers (Fig. 5 and Fig. 6). The linear regression coefficient values are the lowest for trained singers group. It is possible to differ this group from others at a 5% significance level (p = 0.04 for Kruskal–Wallis test).

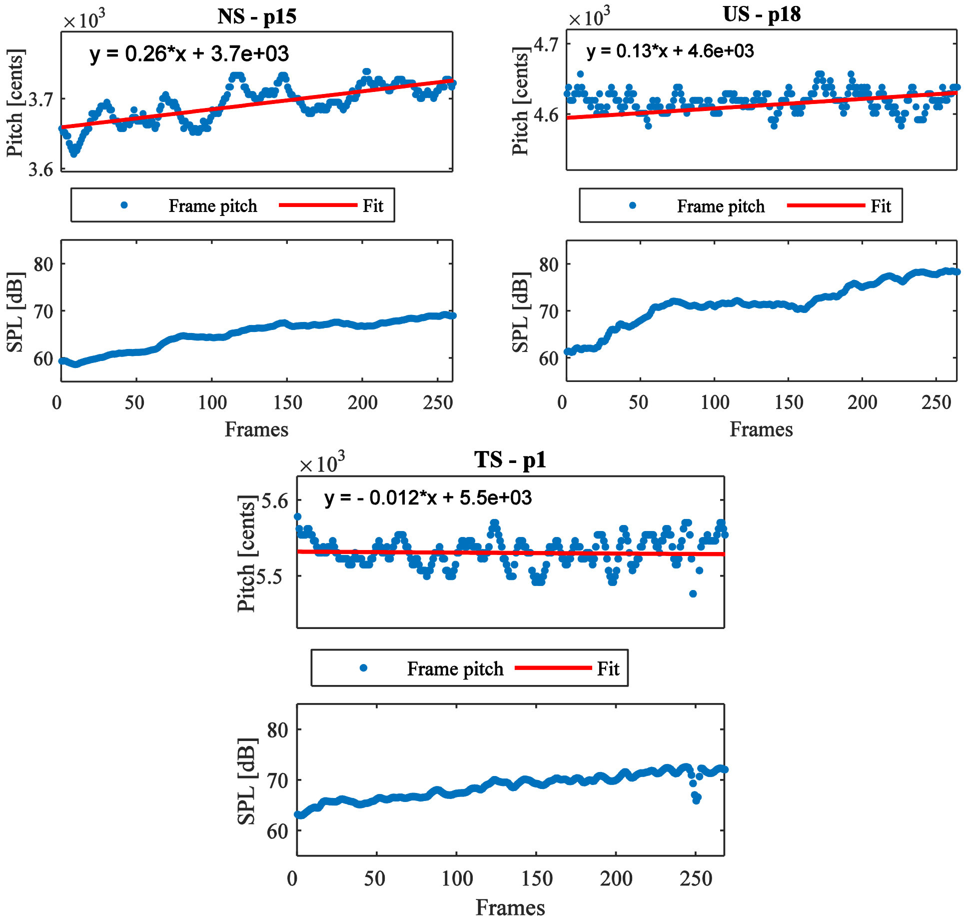

Figure 4. Example of the estimated trend line for nonsinger (NS), untrained (US) and trained (TS) singer subject. The bottom plots present Sound Pressure Level values

4. Discussion

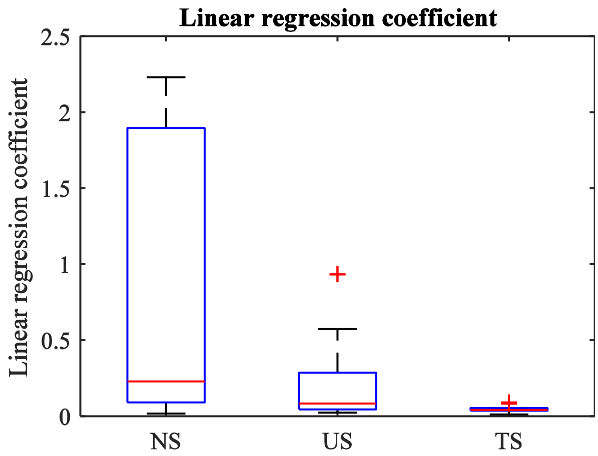

Figure 7. Boxplots median of linear regression coefficients from a vowel during crescendo task. NS – nonsingers, US – untrained singers, TS – trained singers

The wrong intonation for p3, and p9 is connected with the level of training: both singers are at the beginning of the voice lessons process. Additionally, p9 is an untrained singer with a couple of years of experience, whereas p3 is nonsinger. It caused differences in errors (170 cents).

The frequency resolution in all studied groups is lower than one tone (mean < 150 cents) and can be used as the voice advancement indicator (mean NS: 120 cents, US: 55 cents, TS: 20 cents). Mean pitch error at the level of 20 cents for trained singers group is similar to [11] results. The positive correlation between pitch error and jitter has not been noticed in the literature before. This trend results from the fact that pitch error is calculated as the absolute value of difference from the pattern tone. It means that greater frequency fluctuations cause the parameter to increase. Lower jitter values for trained subjects (median NS: 0.59%, US: 0.58%, TS: 0.21%) confirms pitch interval accuracy estimation for poor and good singers from [10]. Low values of Jitter indicate the ability to maintain a fixed frequency throughout the singing when it is needed.

The highest, positive regression coefficients for nonsingers during crescendo exercise proved the tendency to increase tone together with Sound Pressure Level as [3]. Additionally, this parameter can be used with others to distinguish trained from untrained singers (median US: 0.06, TS: 0.001), which corresponds to standard values from the literature [11]. Vibrato was not observed in the nonsingers group, whereas it was more regularly for trained than untrained singers. This behavior is well known in the literature [19], but it is worth to sign that it can be useful for the determination level of voice training and advancement by the regularity factor.

Table 2. Statistical parametrization for pitch error.
NS – nonsingers, US – untrained singers, TS – trained singers

Vowel	NS [cents]			US [cents]			TS [cents]		
	Median	Min	Max	Median	Min	Max	Median	Min	Max
a	22.93	144,99	27.66	40.42	64.75	26.76	10.36	17.23	21.73
e	18.89	50.35	28.01	40.93	64.67	30.04	15.78	21.48	21.93
i	28.33	154.86	27.79	23.42	47.75	22.86	16.31	23.81	31.39
o	28.25	144.25	24.79	32.85	56.51	26.76	15.47	20.13	24.03
u	28.22	109.44	167.29	21.87	46.23	53.92	23.42	30.04	33.81

Table 3. Linear regression analysis results from crescendo sounds.
NS – nonsingers, US – untrained singers, TS – trained singers

NS			US			TS		
Median	Min	Max	Median	Min	Max	Median	Min	Max
0.160	0.123	0.615	0.064	0.137	0.279	0.001	0.011	0.019

5. Conclusions

We confirmed that there are parameters like pitch error, jitter, linear regression coefficients, and vibrato that differ significantly among nonsingers, untrained and trained singers. Low values of linear regression coefficients for trained and untrained voices indicates the ability to keep the pitch constant during a crescendo, which is the pitch stability and the voice quality indicator.

Choosing the promising parameters is the first step to create singers voice quality system. In the future, we will be looking for other ones to perform the classification of the advanced level of singing by, e.g., HMM, ANN, SVM.

References

1. L. Enflo, J. Sundberg, A. McAllister, *Collision and Phonation Threshold Pressures Before and After Loud, Prolonged Vocalization in Trained and Untrained Voices*, Journal of Voice, **27** (September 2013) 527 – 530.
2. I. R. Titze, *Phonation threshold pressure: a missing link in glottal aerodynamics*, The Journal of the Acoustical Society of America, **91** (May 1992) 2926 – 35.
3. I. R. Titze, J. Sundberg, *Vocal intensity in speakers and singers*, The Journal of the Acoustical Society of America, **91** (May 1992) 2936 – 2946.
4. J. Sundberg, *Flow Glottogram and Subglottal Pressure Relationship in Singers and Untrained Voices*, Journal of Voice, **32** (January 2018) 23 – 31.
5. N. Henrich, B. Roubeau, M. Castellengo, *On the use of electroglottography for characterisation of the laryngeal mechanisms*, Proceedings of the Stockholm Music Acoustics Conference, August 6–9, 2003.
6. N. Henrich, C. D'Alessandro, B. Doval, M. Castellengo, *Glottal open quotient in singing: measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency*, The Journal of the Acoustical Society of America, **117** (March 2005) 1417 – 30.

7. E. Półrolniczak, M. Kramarczyk, *Formant analysis in the assessment of the quality of choral singers*, Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), (September 2013) 200 – 204.
8. J. Sundberg, *Articulatory interpretation of the "singing formant"*, The Journal of the Acoustical Society of America, **55** (April 1974) 838 – 844.
9. T. Murry, *Pitch–matching accuracy in singers and nonsingers*, Journal of Voice, **4** (January 1990) 317 – 321.
10. T. Nakano, M. Goto, Y. Hiraga, *An automatic singing skill evaluation method for unknown melodies using pitch interval accuracy and vibrato features*, INTERSPEECH 2006, April 2006.
11. E. Półrolniczak, M. Łazoryszczak, *Quality assessment of intonation of choir singers using f0 and trend lines for singing sequence*, Metody Informatyki Stosowanej, **4** (2011) 259 – 268.
12. J. Sundberg, *Intonation in singing*, The Oxford Handbook of Singing, chapter 13, April 2019.
13. S. Ternstrom, J. Sundberg, *Intonation precision of choir singers*, Journal of the Acoustical Society of America, **84** 59 – 69.
14. A. Grell, J. Sundberg, S, Ternstrom, M. Ptok, E. Altenmuller, *Rapid pitch correction in choir singers*, Journal of the Acoustical Society of America, **126**(1) (2009) 407 – 413.
15. J. Dai, S. Dixon, *Intonation trajectories within tones in unaccompanied soprano, alto, tenor, bass quartet singing*, Journal of the Acoustical Society of America, **146** (August 2019).
16. A. Vurma, J. Ross, *Production and perception of musical intervals*, Music Perception, **23** 331 – 344.
17. P. Gramming, J. Sundberg, S. Ternstom, R. Leanderon, W. Perkins, *Relationship between changes in voice pitch and loudness*, Journal of Voice, (1988) 118 – 126.
18. K. Deepak, *Epoch extraction using zero band filtering from speech signal*, Circuits Systems and Signal Processing, **34** (2014).
19. E. Prame, *Measurements of the vibrato rate of ten singers*, The Journal of the Acoustical Society of America, **96** (October 1994) 1979 – 1984.