



Audio Engineering Society

Convention Paper 10436

Presented at the 149th Convention
Online, 2020 October 27–30

This Convention paper was selected based on a submitted abstract and 750-word precis that have been peer reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This convention paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library, <http://www.aes.org/e-lib>. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Effects of Software Tuning Programs on Vocal Recordings

Chandler Raymond Bridges, Jr.

Jacobs School of Music, Indiana University Bloomington, 107 S Indiana Ave, Bloomington, IN, 47405

Correspondence should be addressed to Chandler Bridges (chanbrid@iu.edu)

ABSTRACT

The primary purpose of this study was to investigate the effect of tuning software on pitch correction for a vocal example. A professional tenor singing the melody from the chorus section of a popular song. Three commonly used pitch-shifting tuning programs to create experimental stimuli. All tuned examples were created with the programs' automatic tuning settings. The examples were analyzed with Praat software to determine the amount of pitch change and the remaining distance to the desired target pitch for each programs' resulting output. One program provided more aggressively tuned results and had closest condition to the desired target pitch; however, there is much debate over which software program is preferred.

1 Introduction

Antares Auto-Tune, a pitch correction effect, has grown in popularity and use in productions of modern recordings since the late 1990s. While many critics cite its use as unfair or even unethical, its use in musical applications remains ubiquitous in modern music production.

Auto-Tune is a software-based music production effect created by Andy Hildebrand in 1997 (Hildebrand, 1999; White, 1999) based on his research and development of a seismic data interpretation workstation he created while working for Exxon Oil Company. With this device, an audio signal, when routed through the Auto-Tune software, is analyzed and the pitch of the signal is modified through a complex set of user controls. Pitch correction is often used to take an out of tune performance, analyze its pitch content, and process the signal to adjust its pitch to the desired frequency. Often, this device is used on a vocal performance; however, it is equally effective in

modifying the pitch of any monophonic solo instrument (Hildebrand, 1999).

In the time since Hildebrand invented Auto-Tune and other musical processors, many competing pitch-correction programs have been released with similar functionality, each with its own benefits and strengths. The most popular of these competitors are Melodyne and WavesTune. These three software programs are the most commonly used for pitch correction in digital music production (Muckala, 2013). Many studies are also using these pitch-correction software programs for control purposes in research experiments (Geringer, 2015; Geringer, 2014; Palispis, 2013; Sundberg, 2013; Presto, 2011; Walsh, 2009).

When a musician or vocalist conveys emotion, feeling, meaning, and can touch the listener in a profound way, the fact that a note or syllable is pitch corrected should make little to no difference to the listener. Due to studio cost considerations, the vocal warmth and strength of the singer at the time of recording, and the day-to-day variations in the operation of human vocal cords, to name just a few common barriers, it is impossible and unrealistic to

expect a perfect performance on every level of judgment. It is for this reason tuning is an indispensable tool in the studio. Tuning programs, like special effects in a movie, can be used sparingly or to excess. In 1998, producers of the Cher hit single “I Believe” employed Auto-Tune with extreme settings to create a never before heard electronic effect on the lead vocal. Quite possibly, it was this usage that brought it into the public scrutiny and initiated the ever since heated debate about its use in music production. However, Auto-Tune was never designed with this intent in mind. Hildebrand created it to be invisible and undetectable when used as he prescribed. Hildebrand has since admitted he never thought anyone would utilize it in the manner it was used on Cher’s vocal.

Tuning software generally has multiple settings to tailor the processing to the specific vocal being treated. All three programs investigated have an automatic tuning mode for use with no user input needed. Not only can be the programs utilized in automatic mode, but also, they offer graphical mode, for more precise control, where the engineer manually adjusts the pitch on an individual note-by-note basis. This graphical process is more precise in that the engineer can decide manually how close or far from the pitch the tuned result would be. Frequently, for a lead vocal or any important part, tuning software is used on a note-by-note basis in a graphical mode (McJunkin, 2004).

Most commercially released recordings are for entertainment. If the recording’s goal is to provide enjoyment, then the use of tuning software can help eliminate distracting or unwanted tuning issues in the performance that have been suggested to interfere with the potential of the emotional impact. More importantly, the true vision of the composer can be realized through the use of modern production techniques of editing and tuning.

Digital audio workstations and tuning software have opened up many new possibilities for the production of musical compositions and are shaping the way music in the 21st century is being created. While there are many published comments and debates regarding which program is most effective (Garner,

2014; Muckala 2013), no study to date has conducted a systematic analysis of the actual results from the processing algorithms used by tuning software. It is the aim of this study to empirically examine the results of the three most commonly used programs in music production: Antares Auto-Tune, Celemony Melodyne, and WavesTune.

2 Method

Analyses were completed using Praat software to analyze each note of the vocal selections. Measurements of average pitch for each note, distance from desired target pitch and overall center of gravity were extracted for each of the four conditions.

2.1 Preparation of Stimuli

We recorded a baritone soloist with 16 years of professional experience as a soloist and in choral ensembles. The singer studied applied voice during undergraduate studies and had taught private voice lessons for 4 years at the time of the recording. We asked him to sing a vocal selection from the chorus of a popular song.

The singer was recorded in a professional recording studio in Los Angeles using a Gauge Precision Instruments ECM-87 cardioid microphone without low frequency roll-off or input padding. He sang at a distance of 12” from the microphone equipped with a Gauge Precision Instruments pop-filter. An ART microphone preamplifier fed the signal to a Pro Tools HD (v.10.2.9), Digital Audio Workstation (AVID HD I/O interface.) Record settings were 24-bit depth at a 44.1kHz sample rate. We processed the recording with three commonly used professional tuning software packages: WavesTune (v.9.1.0.9), Antares Auto-Tune (v. 7.0.9), and Melodyne Singletrack (v. 2.0.) We normalized each processed file to account for subtle decibel level changes post-processing.

2.2 Design and Procedures

In order to investigate the research questions of interest, we created four conditions: the original performance, tuned with Antares AutoTune, tuned with Celemony Melodyne, and tuned with Waves Tune. This resulted in 4 experimental examples.

Each example was then imported into Praat software for analysis. We calculated values for pitch by selecting each note individually in the performance and extracting the value from Praat into an Excel spreadsheet. There were 22 examples selected from each of the four conditions. Additionally, the desired target frequency was calculated for each note performed (Table 1). This value was used to determine the remaining distance, post-tuning, to the exact frequency.

Table 1
Desired Target Pitches

Text	Note	Target Pitch
And	A#3	233.08
these	A#3	233.08
mem	G4	392
or	F4	349.232
ies	D4	293.664
are	F4	349.232
all	F4	349.232
that	D4	293.664
I	G4	392
hoped	F4	349.232
they would	C4	261.632
be	A#3	233.08
with	A#3	233.08
day	G4	392
ay	D4	293.664
light	D4	293.664
they	F4	349.232
might	F4	349.232
just	D4	293.664
fade	G4	392
a	F4	349.232
way	D4	293.664

3 Results

Raw data consisted of pitch analysis of intonation accuracy with Praat software. Table 2 shows composite means and standard deviations of the software output. We screened the data and verified that assumptions of the analysis of variance were met. We employed a one-way analysis of variance (ANOVA) to analyze amount of pitch correction from each program. There was not a significant difference in frequency values from the pitch correction processing, $F(2,42) = 2.06, p > .05$ (Table 3).

Table 2
Means and Standard Deviation (SD) for Amount of Pitch Correction

	Auto-Tune	Melodyne	WavesTune
Mean Hz	19.05	20.60	20.25
SD	7.39	10.24	9.29

Table 3
Analysis of Variance for Amount of Pitch Change (Distance to Target)

Source of Variance	SS	df	MS	F	p
Treatment	29.15	2	14.58	2.06	0.14
Error	296.77	42	7.06		
Ss/Bl	4866.47	21			
TOTAL	5192.40	65			

Figure 1 illustrates the differential effects of pitch corrections for each program measured in Hertz. It can be seen that Celemony Melodyne had the greatest amount of correction ($M = 20.6$ Hz). WavesTune had the next largest amount of correction ($M = 20.25$ Hz) and AutoTune with the least ($M = 19.05$ Hz).

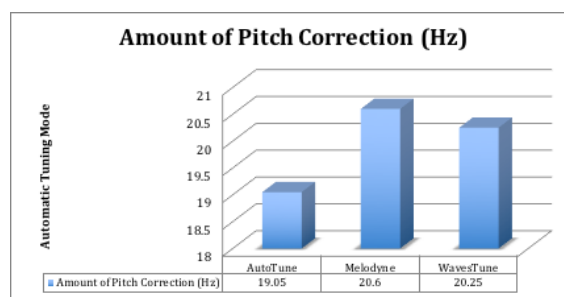


Figure 1. Means for amount of pitch correction (Hz)

Similarly, as shown in Figure 2, the remaining distance to the desired target pitch also indicates Celemony Melodyne with the least remaining out of tune notes ($M = 1.95$ Hz), followed by WavesTune ($M = 2.3$ Hz), and Auto-Tune ($M = 3.5$ Hz).

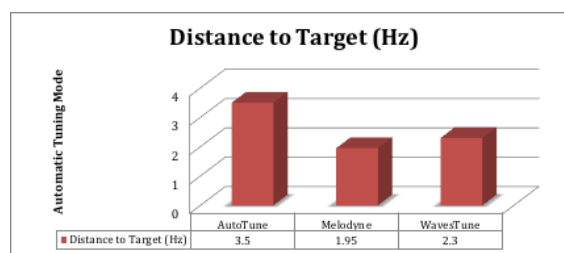


Figure 2. Means for distance from target pitch after processing (Hz)

Figure 3 represents this distance to the desired target pitch in cents: Celemony Melodyne ($M = 9.33$ cents), WavesTune ($M = 10.9$ cents), and Auto-Tune ($M = 16.75$ cents).

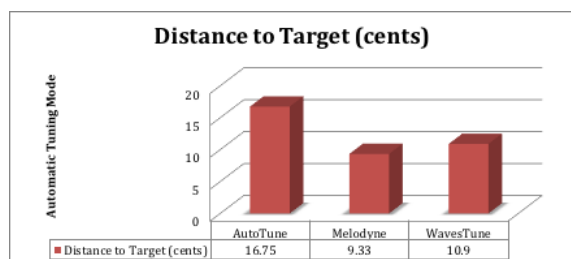


Figure 3. Means for distant to target (cents).

4 Discussion

The results from the Praat analysis found very little frequency difference in processing outcome of this vocal example. While Melodyne did have the most accurate results, there was not a statistically significant difference between the programs. However, as tuning accuracy has been shown to be important, the processing from Melodyne came closest to the desired pitch results and is the most accurate tuning program of the three examined.

This investigation used the pitch correction software in automatic tuning mode. In most professional vocal productions, a more refined “graphical” mode is deployed to correct out of tune performances. It was outside the scope of this article to use this method, as tuning in this method requires subjective judgments on the method, amount, and usage of the tuning tools provided within each software user interface. However, using this “graphical” method of tuning could be systematically applied to the stimuli for additional analyses of the effect of the tuning software in future studies.

It is important to note that only one voice type was examined for this study. A fuller documentation of these tuning programs could certainly be undertaken that expands upon this research, which could include different voice types, gender, and perception-based components, wherein listeners could judge their perceptions of the affected sounds. Additionally, further investigation could surely be done of the variation between degrees of pitch correction applied, percentage of pitch correction applied, and on the reactions of these tuning programs’ algorithms to intonation as a pitch varies more widely from its intended target. Further considerations, such as starting central pitch, 440 Hz

or 442 Hz as an example, could also be of value. This kind of research could easily expand outward even further, to applications of this technology on audibility, the effects of tuning on timbre, and more.

Additionally, often more experienced musicians have been measured to perform with slightly sharp deviation in many studies (Geringer, 1978; Geringer, 1987; Karrick, 1998; Damon, 2014). Future study could investigate vocal performances tuned sharp in relation to the accompanying instrumental tracks for significant preference by listeners.

Lastly, future studies could also compare the tuning results to listener rated preferences for intonation and tone quality. While, correct tuning is important, as shown in this study there is no significant difference in the outcomes between these programs. Therefore, the resulting tone quality of the processed output must be considered in the overall usability and desirability of the vocal selection. Perhaps, this change in tone quality will prove even more important than the amount of vocal correction instantiated by the processing itself.

References

- [1] H. Hildebrand, "Pitch detection and intonation correction apparatus and method." [US patent 5973252](#), assigned to Auburn Audio Technologies, Inc. (1999).
- [2] P. White, "Antares system Autotune: Digidesign Pro Tools TDM pitch correction plug-in." *Sound on Sound* (1997).
- [3] J.C. Muckala, (2013). "Commercial Music: What Every SLP Should Know." *SIG 3 Perspectives on Voice and Voice Disorders*, vol. 23, no. 1, pp. 28-35 (2013).
- [4] J.M. Geringer, C. Bridges, "Perceptual Salience of Amplitude and Frequency Modulation in Vocal Vibrato," presented, at the *June 2015 Voice Foundation Annual Symposium (International)*(2015).
- [5] J.M. Geringer, L.M. Mann, C.R. Bridges, "Effects of Vibrato Extent on Perception of Singer Intonation," presented at the *June 2014 Voice Foundation Annual Symposium (International)*(2014).

- [6] B. Palispis, *A process of producing a record in the emerging digital age*. (1552690, California State University, Los Angeles). ProQuest Dissertations and Theses, p. 88. Retrieved from <http://search.proquest.com/docview/1506154936?accountid=4840>. (1506154936) (2013).
- [7] J. Sundberg, "Acoustic and psychoacoustic aspects of vocal vibrato." Dept. for Speech, Music and Hearing: Quarterly progress (2013)
- [8] P. Presto, "Cosmetic Surgery for the Voice: All's Fair in Love and Pitch." *Journal of Singing*, vol. 67, no. 3, p. 327 (2011).
- [9] C. Walsh, "Pitch-correcting software: Fix it in the DAW?" *Pro Sound News*, pp. 31-32. Retrieved from <http://search.proquest.com/docview/200221334?accountid=4840> (2009).
- [10] J. McJunkin, "Cut and create perfect vocals." *ReMix*, vol. 9, pp. 53-58 (2004).
- [11] K.K. Garner, "Vocal Recording Techniques for the Modern Digital Studio" (2014).
- [12] Praat [Computer software]. Amsterdam, The Netherlands: Institute of Phoenetic Sciences, University of Amsterdam (Version 5.3.58) (2013).
- [13] AVID Pro Tools HD [Computer software]. United States: Pro Tools (Version 10.2.9) (2013).
- [14] AVID HD I/O [Audio hardware]. United States: AVID Audio Interface (2013).
- [15] WavesTune [Computer software]. United States: Waves (Version 9) (2013).
- [16] Auto-Tune [Computer software]. United States: Antares (Version 7.0.9) (2013).
- [17] Melodyne singletrack [Computer software]. Germany: Celemony (Version 2.1.0) (2013).
- [18] J.M. Geringer, "Intonational performance and perception of ascending scales." *Journal of Research in Music Education*, vol. 26, pp. 32-40 (1978).
- [19] J.M. Geringer, C.K. Madsen, "Pitch and tempo preferences in recorded popular music." *In Applications of Research in Music Behavior* (pp. 204-212). Tuscaloosa, AL: University of Alabama Press. 7 (1987).
- [20] B. Karrick, "An examination of the intonation tendencies of wind instrumentalists based on their performance of selected harmonic musical intervals." *Journal of research in music education*, vol. 46, no. 1, pp. 112-127 (1998).
- [21] B. Damon, *Instruction type: Effects on pitch accuracy in female collegiate declared voice majors*. (3618372, Liberty University). ProQuest Dissertations and Theses, pp. 146. Retrieved from <http://search.proquest.com/docview/1528594763?accountid=4840>. (1528594763) (2014).