What About Extent?: **Examining Current Vibrato Metrics**

Theodora Nestorova, M.M., and Joshua D. Glasner, M.M., Ph.D.

Pan American Vocology Association InFormant *Newsletter; Vol. 2, Issue 3; 2021*

***What About Extent?*: Examining Current Vibrato Metrics**

Studies involving vibrato generally reference vibrato rate and extent measurements from previous literature as normative comparisons. Many researchers use these metrics to quantify the complex phenomenon of vibrato, but conventional applications often produce differing results for identical audio samples. However, current methods of measuring vibrato rate and extent may not be consistent across platforms and software programs. Furthermore, only presenting average vibrato rate and extent may not fully characterize non-uniform vibrato. Such discrepancies in the literature may result in unreliable and non-generalizable conclusions.

Sundberg (1987) defines vibrato as a long-term "periodic modulation of the fundamental frequency." The element of time in this definition is crucial, which is a motivating factor in studying the time-varying characteristics of vibrato. Re-examining normative vibrato metrics supports the development of novel analytical methods that consider temporal vibrato variation. Historically, vibrato has been analyzed with tools presuming a Western Classical opera aesthetic, applicable only if vibrato is uniform, consistent, and persistent over time. Analyzing the time-varying aspects of vibrato may promote a broader representation of singing genres with stylistically distinct vibrato features, and should be the subject of future inquiry.

In the Western Classical singing tradition, vibrato is typically periodic and sinusoidal, so its properties have been quantified using averages. Vibrato rate refers to the number of fundamental frequency (f_o) oscillations per second, and is usually measured peak to peak. Typically, vibrato extent describes how far above and below the mean f_o the frequency oscillates each cycle. Extent is usually expressed as either full or half, the former a measure of peak to trough and the latter a measure of peak to a calculated mean f_o. Extent is reported as either a percentage of f_o or in cents, with 6% or 100 cents making one semitone (in equal temperament).

Beyond rate and extent, other measurable acoustic characteristics of vibrato include vibrato jitter and vibrato shimmer, measures of perturbations in the f_o and amplitude, respectively, and onset/offset modulation, describing the quality of the first and last several cycles of the vibrato. Recent investigations have also highlighted the interactions of airflow and intensity with frequency vibrato (Horii, 1989; Nandamudi, 2017; Nandamudi & Scherer, 2019).

Comparison of Vibrato Metrics in Current Voice Analysis Platforms

Throughout the 20th and 21st centuries, applications for spectrographic analysis have entered the mainstream of voice science and pedagogy. Many of these extract vibrato parameters as useful, practical metrics. Yet, a comprehensive comparison of vibrato metrics reported in the currently available voice analysis software does not exist.

Furthermore, several recent investigations examine vibrato differences from a historical perspective (Rothman, Diaz, & Vincent, 2000; Ferrante, 2011; Crutchfield, 2012; Howell, 2015; Glasner & Johnson, 2020) using modern analysis techniques. In these studies, there is considerable variability in the reported vibrato measurements of historical versus modern-day recordings. This variability may be caused by the algorithms of each respective program used for analysis. Therefore, it is critical to study further the commonly accepted algorithms used in various programs to evaluate consistency of results across various platforms for vibrato analysis.

To carry out a comparative evaluation, we used three samples (Tokens 1-3) from a professional operatic soprano singing a *messa di voce* on the pitch C_4 and analyzed them with five algorithms, *VoceVista 3.3.7, VoceVista 3.4.3b*, *VoceVista Video Pro*, a custom *Praat + R* code by Nestorova, Howell, & Gilbert (2021), and a custom *Python* script from Herbst, Hertegard, Zangger-Borch, & Lindestad (2017).

Table 1
Resulting fundamental frequency (f_o), vibrato rate (Hz), and vibrato extent

measurements and averages from 5 algorithms. Measurements taken from 6 complete cycles in the middle of a messa di voce *task sung by a professional operatic soprano.*

Algorithm	**Sample**	**Reported Metrics**		
		***f_o* (Hz)**	**Rate (Hz)**	**Extent (cents)**
VoceVista 3.3.7	Token 1	538	6.3	101
	Token 2	536	6.4	109
	Token 3	531	6.3	106
	Token Avg	535	6.3	105
VoceVista 3.4.3b[1]	Token 1	537.8	6.4	169
	Token 2	535.5	6.5	180
	Token 3	530.6	6.4	194
	Token Avg	534.6	6.4	181
VoceVista Video Pro	Token 1	535.2	6.3	223
	Token 2	532.7	6.5	231
	Token 3	528.8	6.3	235
	Token Avg	532.2	6.4	230
Praat + RStudio Code	Token 1	536.4	6.3	208
	Token 2	532.1	6.4	223
	Token 3	527.4	6.3	227
	Token Avg	532.0	6.3	219
Python Code[2]		--	**Rate (Hz)**	**$\overline{\Delta c}$ (cents)**
	Token 1	--	6.28	67.76
	Token 2	--	6.51	75.18

[1] "According to Miller & Schutte (Voce Vista developers), a minimum of three or four vibrato cycles are required to measure vibrato rate and extent, and at least 10 vibrato cycles are required to measure vibrato jitter accurately." (Guzman, et al., 2012).

[2] See Herbst, et al. (2017) for full description of algorithms and metrics used for vibrato analysis. It should be noted that $\overline{\Delta c}$ there is not an extent or half extent estimate, but is "the average absolute deviation from the mean musical pitch Dc."

Token 3	--	6.40	76.52
Token Avg	--	--	--

Observed Discrepancies Across Algorithms

While f_o and vibrato estimates remained relatively consistent, vibrato extent values varied across all five platforms. This discrepancy certainly deserves attention as it creates challenges for comparing reported vibrato extent measurements in existing literature.

Indeed, the results of this preliminary investigation revealed that comparing vibrato extent may not be straightforward across studies. First, there is the terminological issue; one must understand whether vibrato extent (as is common) refers to full vibrato extent— the amplitude between the peak and the trough of the vibrato wave, or the half extent— the absolute deviation from the mean fundamental frequency. Second, and perhaps more complex, is the question of documentation in reporting the nonuniformity of real-life f_o signals that may deviate from an ideal sinusoid.

From the data reported above, it seems likely that the extent values reported in both *VoceVista 3.3.7* are, in fact, half extent values rather than full. Simply doubling these values results in extent measurements that are nearly identical to the *Praat + R* code used in this preliminary comparison (Nestorova, et al., 2021).

It is important to note that an earlier version of *VoceVista 3.4.3* included algorithmic inaccuracies that affected vibrato measurements. The creators of the software have fully disclosed this and provided an updated, version of the software (*VoceVista 3.4.3b*). It must also be acknowledged that the differences in reported values across platforms may be attributed to variations in the discretization and the time-step function. The significance of this factor in vibrato metric calculations warrants further investigation.

Existing Half Extent Parametric Models in Previous Literature

The singing voice science community has not reached a consensus about the analysis of vibrato extent. Historically, the metric has been expressed in semitones, cents, and percentages, reporting these units somewhat inconsistently (both with +/- semitones and Hertz). While vibrato full extent is the average result of the differences of two successive peaks and troughs of a sinusoidal wave, half extent is the excursion value from a calculated mean f_o to top peak and from that same reference to bottom trough. While further study is still necessary to reach a conclusion, reporting half extent rather than a full extent expresses extent in a way that is readily interpretable for detailed analysis. The half extent procedure provides a solution for asymmetrical measurements in reference to vibrato, since the upper extent (peak to midline) may differ from the lower extent (trough to midline) in an imperfect, uneven sinusoid.

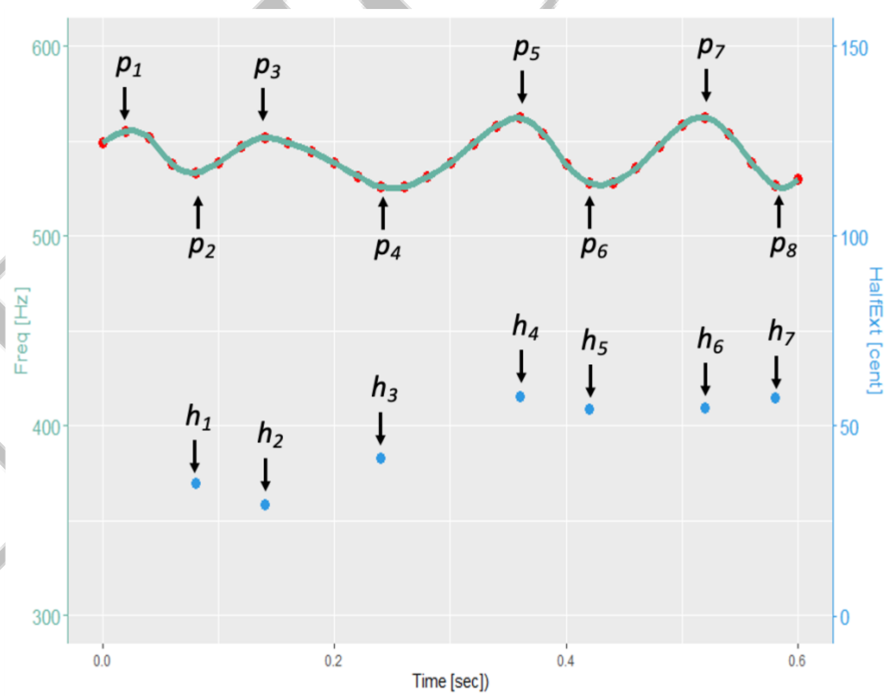

Figure A1. Illustration of the half extent calculation from the measured f_o contour.

Legend: Blue curve - f_o contour; Red dots – measured $\boldsymbol{f_o}$ contour; $\boldsymbol{p_i}$ points – peaks (p_1, p_3, p_5, and p_7) and troughs (p_2, p_4, p_6, p_8); blue $\boldsymbol{h_i}$ points – calculated half extents in cents

Algorithm

1. Convert the measured f_o contour into cents (formula found in Herbst, 2012).
2. Identify peaks (p_1, p_3, p_5, and p_7) and troughs (p_2, p_4, p_6, p_8).
3. For each successive each peak-trough pair, calculate full extent:
   $e_i = | p_{i+1} - p_i |$
4. Calculate half extent:
   $h_i = 0.5 * e_i$

A full extent is then subsequently calculated by taking the sum of each resulting half extent value and dividing by the total number of half extent pairs. Thus, a more credible extent measurement considers a distinction in half excursion values above and below a reference pitch, and is gathered in its half and full excursion versions.

Further Extent Considerations

This preliminary study reveals inconsistencies in vibrato extent values across platforms. This highlights the ambiguity surrounding algorithmically deriving and reporting vibrato extent in current and previous singing voice studies. Although such a granular investigation may seem insignificant or even “nit-picky,” discrepancies as highlighted in this report may be problematic for researchers interested in contextualizing their findings with past literature. As such, this preliminary inquiry justifies the need for transparency, specificity, and reassessment when reporting vibrato extent values in future research.

One possible solution to the issue of variance in reported vibrato extent is for investigators to report vibrato extent results by indicating both the mathematical methodology for calculation and, by including +/- before the vibrato extent value using half extent measures to indicate this measurement. Qualifying and discerning half from full vibrato extent measures in both singing voice research methodology and analysis is essential.

While the fields of singing voice science and voice pedagogy have grown significantly since the mid-20th century, it seems likely that a foundational measurement— vibrato extent— in our field has been studied and reported using inconsistent methods. Understanding each algorithm's method for calculating singing voice measurements such as vibrato is paramount for accurately reporting normative data. It is our hope that this report highlights the discrepancies found within the literature about the measurement and analysis of vibrato in singing voices, and that it helps the field work toward future consensus and better understanding of vibrato in the singing voice.

Acknowledgements:

We would like to thank Dr. Christian Herbst for his work analyzing the audio samples with his Python vibrato extraction algorithm and for sharing helpful graphical outputs with us during the drafting process. Additionally, we would like to thank Dr. Ian Howell and Professor Josh Gilbert for many helpful conversations related to this topic as well as for their contribution in developing the algorithm that was used in the R Studio portion of this analysis.

References

Boersma, P., & Weenink, D. (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.50. Retrieved from http://www.praat.org/

Crutchfield, W. (2012). Vocal performance in the nineteenth century: An overview. In C. Lawson, & R. Stowell (Eds.), *The Cambridge History of Musical Performance*. New York, NY: Cambridge University Press.

Ferrante, I. (2011). Vibrato rate and extent in soprano voice: A survey on one century of singing. *The Journal of the Acoustical Society of America, 130*(3), 1683-1688. https://doi.org/10.1121/1.3621017

Glasner, J.D., & Johnson, A.M. (in press). Effects of Historical Recording Technology on Vibrato in Modern-Day Opera Singers. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2020.07.022

Herbst, C.T. (2012). Freddie Mercury - Acoustical voice analysis. *L.O.G.O.S. Interdisziplinair.* 20. 174-183. www.doi.org/10.7345/prolog-1203174

Herbst, C.T., Hertegard, S., Zangger-Borch, D., & Lindestad, P. (2017). Freddie Mercury—acoustic analysis of speaking fundamental frequency, vibrato, and subharmonics. *Logopedics Phoniatrics Vocology*, 42(1). https://doi.org/10.3109/14015439.2016.1156737

Horii, Y. (1989). Acoustic analysis of vocal vibrato: A theoretical interpretation of data. *Journal of Voice,* 3(1), 36-43. https://doi.org/10.1016/S0892-1997(89)80120-1

Howell, I. (2015). Unmasking Dame Nellie Melba's Vibrato: The Challenge of Drawing Technical Conclusions from Historical Recordings. Retrieved from http://vocped.ianhowell.net/melba/

McCoy, S. (2019). *Your Voice: An Inside View 3rd Edition.* Inside View Press, Delaware, 91.

Miller, D. G., Horne, R., & Schutte, H. (2008), *VOCEVISTA (visual feedback for instruction in singing)*. 34.

Nandamudi, S. (2017). *Aerodynamics of Vocal Vibrato* (Publication No. bgsu1499427478103556) [Doctoral dissertation, Bowling Green State University]. Retrieved from http://rave.ohiolink.edu/etdc/view?acc_num=bgsu1499427478103556

Nandamudi, S., & Scherer, R.C., (2019), Airflow Vibrato: Dependence on Pitch and Loudness. *Journal of Voice*, 33(6), 815-830. https://doi.org/10.1016/j.jvoice.2018.05.007

Nestorova, T.I.. (2021). *Does Vibrato Define Genre or Vice Versa?: A Novel Parametric Approach to Complex Vibrato Patterns* [Unpublished master's thesis]. New England Conservatory of Music.

Nestorova, T.I., Howell, I., & Gilbert, J. (2021, June). *Analysis and Interpretation of Complex Vibrato Patterns: Deriving Novel Parameters of Variable and Irregular Vocal Vibrato*. Poster session presented at *The Voice Foundation* 50th Anniversary Symposium, Online.

Nix, J., Perna, N., James, K., & Allen, S. (2016). Vibrato rate and extent in college music majors: A multicenter study. *Journal of Voice*, 30(6), 75-e41. https://doi.org/10.1016/j.jvoice.2015.09.006

Prame, E. (1997). Vibrato extent and intonation in professional Western lyric singing. *The Journal of the Acoustical Society of America,* 102: 616-621. https://doi.org/10.1121/1.419735

Rasch, R.A. (1985). Jitter in violin and singing-voice tones. *The Journal of the Acoustical Society of America,* 78. https://doi.org/10.1121/1.2022767

Rothman, H. B., Diaz, J. A., & Vincent, K. E. (2000). Comparing historical and contemporary opera singers with historical and contemporary Jewish cantors. *Journal of Voice,* 14(2), 205-214. https://doi.org/10.1016/S0892-1997(00)80028-4

Shannon, C. E. (1949). Communication in the Presence of Noise. *Proceedings of the IRE*, *37*(1), 10–21.

Sundberg, J. (1987). *The Science of the Singing Voice*. Northern Illinois University Press, Dekalb, IL, 170.

Sundberg, J. (1994). Acoustic and psychoacoustic aspects of vocal vibrato. *KTH School of Computer Science and Communication Transmission Laboratory* Q Prog Status Rep. 35: 45–68. Retrieved from https://www.speech.kth.se/prod/publications/files/qpsr/1994/1994_35_2-3_045-068.pdf

Sundberg, J. (1975). Vibrato and vowel identification. *KTH School of Computer Science and Communication Transmission Laboratory* Q Prog Status Rep.; 16:2-3: 49-60. Retrieved from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.431.9117&rep=rep1&type=pdf

Titze, I., Story, B., Smith, M., & Long, R. (2002). A reflex resonance model of vocal vibrato. *The Journal of the Acoustical Society of America*, 111: 2272-82. https://doi.org/10.1121/1.1434945

VoceVista 3.3.7. (2018). *VoceVista* (Version 3.3.7) [Computer software app]. Online https://vocevista.software.informer.com/3.3/

VoceVista 3.4.3.b (2019). *VoceVista* (Version 3.4.3b) [Computer software app]. Online https://vocevista.software.informer.com/3.3/

VoceVista Video Pro. (2021). *VoceVista* (Version 3.3.7) [Computer software app]. Online https://www.vocevista.com/support/vocevista-video-support/

Winckel, F. (1953). Physikalische Kriterien für objektive Stimmbeurteilung, *Folia Phoniatrica Logopaedica*, 5: 232-252. PMID: 13142111