Vibrato rate and extent in soprano voice: A survey on one century of singing

Isidoro Ferrante[a)]
Università degli studi di Pisa, Dipartimento di Fisica, largo Bruno Pontecorvo 3, I-56127 Pisa, Italy

(Received 25 March 2011; revised 30 June 2011; accepted 4 July 2011)

This work presents a statistical study of vibrato parameters in soprano voices. More than one hundred recordings of the same tone sung by 75 artists have been analyzed. Vibrato rate and extent, tone length and intonation, together with their correlations are the main parameters under examination. The study shows a clear decrease of the mean vibrato rate during the last century (-1.8 ± 0.3 Hz/century), together with an increase of vibrato extent (56.4 ± 0.3 cent/century). Vibrato rate and extent show a statistically significant negative correlation ($r = -0.62$). Vibrato rate increase near the end of the tone has been observed too, in agreement with previous measurements, together with a mean increase of the pitch of the tone. A small positive correlation has been also found among note duration and vibrato extent. © *2011 Acoustical Society of America*. [DOI: 10.1121/1.3621017]

PACS number(s): 43.75.Rs [DD] Pages: 1683–1688

I. INTRODUCTION

Vibrato is a rapid (4–7 Hz) modulation of the sung pitch which is found in various degrees in singers educated in the western classical tradition. Due to the vocal tract resonances, frequency modulation is usually accompanied by an amplitude modulation as well.[1] The physical aspects of vibrato have been studied for more than 80 years, beginning with the pioneering works of Seashore,[2] who published vibrato parameters measured in recordings and laboratory performances. The work of Seashore, however, although of fundamental importance, is outdated not only due to the primitive technologies used, but also because his measurements do not represent the current situation of modern singing. For example, the vibrato rate measured in famous artists of the day reported a minimum rate of 5.9 Hz; today, however, vibrato rates smaller than 5 Hz are quite common.

Several studies have been performed to establish the findings of Seashore: some of them used laboratory measurement on the singer's body to enlighten physiological aspects, others have instead been performed on studio recordings of world famous singers. While the first ones allow the researchers to gain access to the behavior of the phonatory organs, the second ones have a few advantages: the study material is readily available with little effort, measurement on world class singers can be performed (while in laboratory it is easier to have conservatory students), and last, but not least, one has access to the voices of great singers of the past.

A first example of this technique can be found in the work of Keidar *et al.*,[3] where ten recordings of *Di quella pira* from Verdi's *Il Trovatore* have been analyzed. A more comprehensive study is the one by Prame[4,5] who studied more than two hundred tones sung in ten recordings of Schubert's *Ave Maria*. One of the most interesting features found by Prame is the increase of vibrato rate near the end of the tone, an increase that can be found even in violin players. While Prame's studies are of paramount importance, they cannot be taken as representative of Western classical singing due to the choice of a German lied of limited extensions as the material of study.

Bretos and Sundberg, on the contrary, studied two long sustained notes in ten recordings of *Oh patria mia* from *Aida* by Verdi: they confirm and extend the findings of Prame, but their study suffers from low statistic due to the limited number of tones under investigation.

Modern digital technologies have put a very large sample of recorded material to the disposal of scholars and scientists, and more than one century of singing is now available for study, allowing one to compare modern and historical performances: see, for example, the paper by Rothman *et al.*,[6] where the evolution of cantorial singing is put under examination.

The present work differs somewhat from its predecessors since a single tone for every recorded performance has been analyzed; however, the large number of recordings used allows one to collect significant final statistics, and the large number of singers allows one to have results which are independent of the characteristics of single individuals. Moreover one also has the opportunity to study the variations of taste and performance practice throughout the 20th century.

II. DATA SAMPLE

The tone used for this work is the B-flat on the word "*Signor!*" in the aria "*Vissi d'arte*" from *Tosca* by Giacomo Puccini (see Fig. 1). This aria has been chosen for a variety of reasons: first of all, there is a documented unbroken performing tradition spanning more than one century from the first year after the first performance to the present. Moreover, the aria has been performed by a great variety of soprano voices, from lirico to spintos, with a few leggero and even some mezzos. There are also multiple recordings from the same soprano, in a time span of some tens of years. This allows not only statistically significant results, but also a

[a)]Author to whom correspondence should be addressed. Electronic address: isidoro.ferrante@df.unipi.it

FIG. 1. The two beats from *Vissi d'arte* around the tone used in this work.

historical perspective on how vibrato changed in singing and in the taste of the public. The chosen tone overlaps partly with string accompaniment: however, spectrogram analysis shows that in the time-frequency plane, voices and instruments are clearly separated.

Recordings have been collected from all possible sources with an attempt to cover the period from 1900 to 2010 uniformly: while this was not entirely possible, a minimum of five recordings for every decade have been found. Much of the material was available only in compressed format: this, however, has no influence on vibrato parameters. All the obtained recordings have been accepted, provided that the date of the performance could be established with accuracy (usually within one year, in some doubtful case within 2 or 3 yr) and that at least one harmonic could clearly be seen in the spectrogram. A few recordings were rejected since the analysis algorithm failed in providing a reliable tracking of f_0, even with an accurate trimming of the input parameters.

The final sample consists of 105 tones, sung by 75 singers. The oldest one is a 1901 Russian recording, sung by Medea Mei Figner;[7] the newest one is a live performance from Micaela Carosi (Parma, 2009). The most represented singer is Maria Callas, with eight recordings, followed by Renata Tebaldi with six. A short summary of the recordings used can be seen in Table I.

III. ANALYSIS METHOD

The analysis consisted of the tracking of the fundamental frequency f_0 and in the subsequent study of the f_0 vs time curve obtained. It was performed with MATLAB®[8] scripts.

The tone was isolated from the rest of the aria, and saved in PCM format, with a sampling frequency of 44 100 Hz. Left and right channels in stereo recordings have been added so as to have a mono stream.

The preliminary operation consisted in displaying the spectrogram on the screen and in selecting by hand the time-frequency region spawned by the fundamental harmonic. In the great majority of cases the starting and ending times of the tone could be identified with low incertitude; sometimes, however, there were difficulties, mainly when the singer used portamento.

After that, the data stream was divided into overlapping segments of 1024 samples: each segment was delayed with respect to the previous one of 147 samples, so as to sample f_0 at a frequency of 300 Hz. For each segment, four FFTs

TABLE I. List of artists and year of recordings.

Ader	1928	Dessí	2004	Mattila	1997	Salazar	1997
Agostinelli	1908	Destinn	1914	Mei-Figner	1901	Sass	1979
Alda	1914	Durbin	1943	Melba	1907	Scotto	1980
Arangi-Lombardi	1933	Eames	1905	Melba	1910	Spani	1929
Bampton	1942	Eames	1908	Melis	1929	Stapp	1981
Bartolomasi	1919	Elizza	1909	Michael	2007	Steber	1959
Boninsegna	1908	Farrar	1909	Milanov	1957	Sutherland	1968
Brouwenstijn	1951	Francillo-Kaufmann	1925	Moore	1937	Te-Kanawa	1990
Caballé	1976	Freni	1966	Muzio	1940	Tebaldi	1949
Caballé	1978	Freni	1982	Nilsson	1969	Tebaldi	1951
Caballé	1980	Freni	1985	Novotna	1944	Tebaldi	1954
Callas	1950	Freni	1992	Obratzsova	1980	Tebaldi	1956
Callas	1951	Gheorghiu	2000	Olivero	1940	Tebaldi	1959
Callas	1952	Gheorghiu	2006	Olivero	1957	Tebaldi	1964
Callas	1956	Gordoni	1967	Olivero	1960	Tomowa-Sintow	1988
Callas	1958	Holleque	1989	Olivero	1961	Turner	1926
Callas	1964	Janowitz	1970	Pampanini	1939	Turner	1933
Callas	1965	Jeritza	1914	L. Price	1955	Urmana	2005
Callas	1974	Jones	1974	L. Price	1958	Vallin	1932
Caniglia	1930	Kabaivanska	1973	L. Price	1965	Verrett	1978
Carosi	2009	Kabaivanska	1976	L. Price	1972	Vishnevskaya	1976
Cavalieri	1910	Kabaivanska	1978	L. Price	1988	Välkki	1962
Cedolins	2006	Katsuli	2008	Raisa	1933	Welitsch	1948
Cerquetti	1956	Kirsten	1957	Reining	1941		
Crider	2003	Lehmann	1938	Rethberg	1924		
Dabusti	2005	Lubin	1927	Ricciarelli	1984		
Delli-carri	2009	Marton	1994	Rysanek	1953		

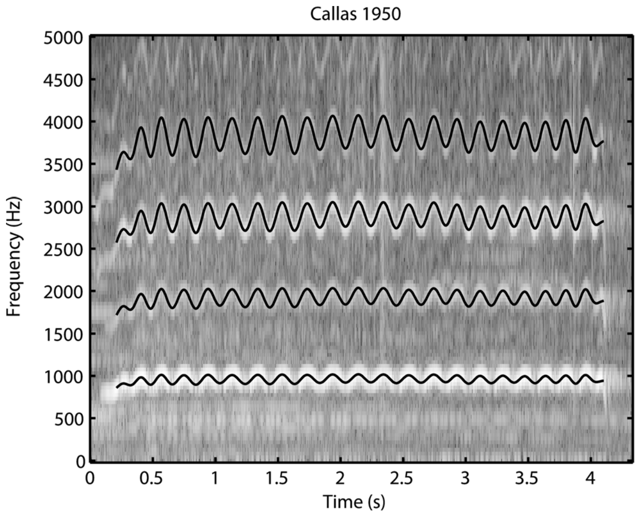

FIG. 2. Example of f_0 reconstruction. The lines corresponding to $2f_0$, $3f_0$, and $4f_0$ are also shown for comparison with higher harmonics.

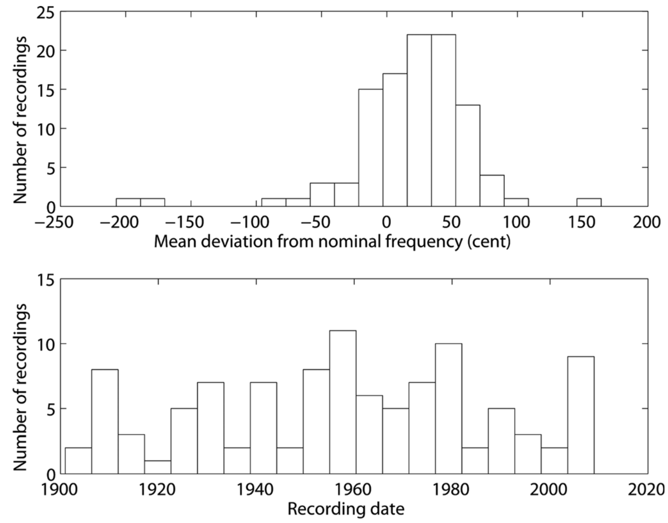

FIG. 3. Top: distribution of $MF0$. A tuning of 440 Hz has been assumed, corresponding to a tone frequency of 932.33 Hz. As can be seen, differences up to two semitones can be seen. It is difficult to know whether these are due to tuning differences, transposition, or to wrong playing speed. See the discussion in the text. Bottom: distribution of recording dates. As can be seen, the distribution is not uniform; however, all decades are well represented.

have been computed with different zero padding: the padding has been chosen so that if f_0 falls in the kth bin for the first transform, the second harmonic falls in the kth bin of the second one, the third harmonic in the kth bin of the third one, and so on. In that way, summing up bin by bin the modules of the various FFTs (properly weighted to account for the different padding length) is equivalent to summing up the four harmonics of every frequency. The f_0 is then chosen as the greatest maximum in the frequency interval selected at the beginning. This technique allows one to automatically select the loudest harmonic and to reach the best resolution obtainable given the data. The resolution obtained on a single point depends on the stronger harmonics, and can vary between 43 and 11 Hz (see Fig. 2).

The resulting f_0 curve is then passed through a low pass–zero phase filter with a cutoff frequency of 10 Hz. The vibrato cycles have then been found by looking at the local maxima. Maxima less than 0.1 s apart can be due only to noise fluctuations: in this case the smallest one has been eliminated. To improve precision, the peak positions t_n have been calculated by interpolating a parabola among the three measured values around the maxima. For each vibrato cycle n, the vibrato rate VR_n has been evaluated as the inverse of the distance between two subsequent maxima. The intonation $F0_n$ is the mean of f_0 between two maxima: it can be measured in Hz, or converted to cent using the nominal tone frequency as reference.

Finally, the vibrato amplitude VA_n has been calculated by taking the rms of f_0 between the two maxima, multiplied by $\sqrt{2}$: this quantity is equal to the amplitude only in the case of a sinusoidal waveform, but in the general case is a good estimate of the fluctuations of f_0 from a constant value.

Correlation among these quantities, and among these quantities and the time t_n have been calculated, excluding, however, the first and last cycles which are not always well measured.

Global quantities have been defined excluding the first and last vibrato cycles: the mean vibrato rate MVR is thus the number of vibrato cycles between the second and last but one maximum divided by the distance among them.

The mean vibrato amplitude MVA is the square root of the mean squared amplitude of all vibrato cycles, excluding the first and the last one.

Finally, the mean intonation MF0 is the mean of f_0 between the first and last but one maximum. MF0 has been measured in cent assuming a standard tuning of 440 Hz.

To allow comparison with previous works, vibrato amplitude has been converted to cent using MF0 as reference, obtaining thus the mean vibrato extent MVE.

Relative resolution on MVR is on the order of 3.3 ms divided by note length; resolution on MF0 is on the order of VA divided by the square root of number of cycles.

The algorithm was applied to the same recordings used by Bretos and Sundberg and gave similar results.

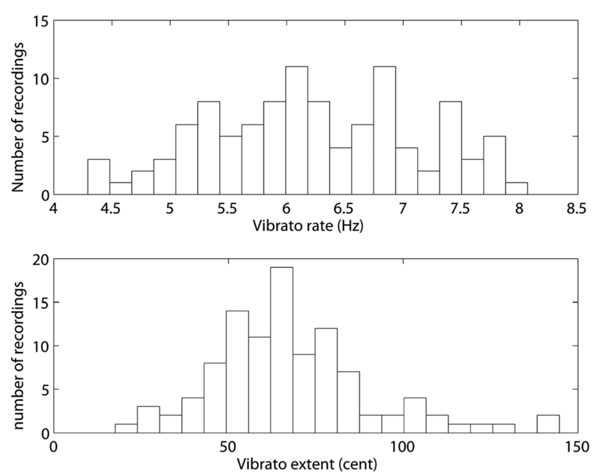

FIG. 4. Distributions of the observed vibrato rate (top) and extent (bottom).

TABLE II. Extreme values measured in the full sample.

Singer	Year	MF0 (cent)	MVR (Hz)	VE (cent)	Duration (s)
Gundula Janowitz	1970	38	8.1	31	3.67
Maria Callas	1964	89	4.2	70	4.5
Mirella Freni	1992	15	5.3	143	4.1
Emma Eames	1905	2.	7.8	17	3.0

IV. RESULTS

The distribution of MF0 (Fig. 3) shows that more than half of the tones are within the nominal frequency of 932.33 Hz plus one quarter of tone; 96% remains within plus or minus one semitone, while in one case (Lehmann) a deviation of plus one semitone and half and in two cases (Turner 1926, Francillo-Kaufmann) minus two semitones have been observed. These deviations are due to several causes:

(1) Orchestral tuning different from the standard.
(2) Off pitch singing.
(3) Transposition.
(4) Wrong recording or playing speed.

Only case (4) can alter the results of this analysis. However, it is not easy to distinguish between this case and cases (1) and (3): while for most modern recordings we can be confident in recording speed, for the older ones and for the ones performed with home equipment we cannot be sure. Moreover, since many of the recordings used come from home-made transfers, also the playing speed suffers from some unknown incertitude, which cannot, however, be bigger than ±12%. Since the effects on vibrato rate are small and on vibrato extent are absent, no effort has been made to correct for this problem.

The mean vibrato rate in the sample is 6.22 Hz, with a standard deviation of 0.90 Hz (Fig. 4). The fastest observed rate is 8.1 Hz in the Gundula Janowitz 1970 recording, while the slowest one is 4.2 Hz in the Maria Callas 1964 recording. The mean vibrato extent is 68 cents with a standard deviation of 24 cents. The greatest measured vibrato extent belongs to Mirella Freni in 1992 (145 cents, almost three quarter of a tone) while the smallest one (only 17 cents) has been measured in the voice of Emma Eames. Vibrato parameters for these extreme values are given in Table II

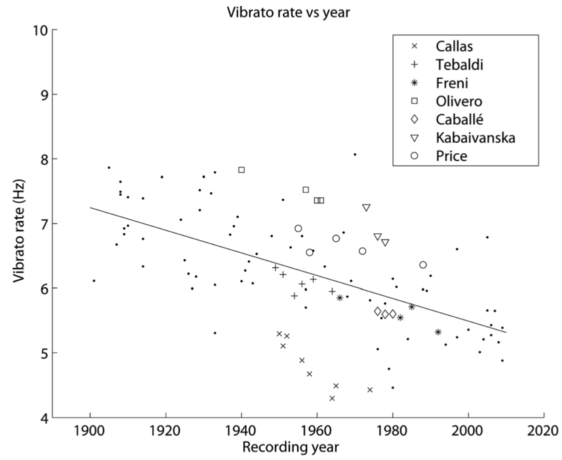

FIG. 6. Vibrato rate as function of the recording year. A clear decrease can be seen. Singers represented in at least three recordings are marked individually: the slowing due to aging is clearly visible.

Figure 5 shows the distribution of the correlation among vibrato parameters and between them and time in the same recording. A clear positive correlation among time and intonation is evident: there is a moderate increase of pitch during tone production. In a similar way, a negative correlation among rate and vibrato extent can also be seen: we will observe it again later in a different context. A smaller similar correlation can be observed between rate and extent and as a consequence also between rate and intonation. As can be seen, instead, the distribution of the correlation among vibrato rate and time is rather well centered around zero: however, an increase of the vibrato rate in the last ten cycles, as described by Prame in Ref. 4, can clearly be observed. Following the procedure described in Ref. 9, the vibrato rate of the last ten cycles has been normalized with the one measured just before, and then averaged over all singer, obtaining thus the normalized vibrato rate NVR. The NVR can be fitted as function of the vibrato cycle k (with $k = [-9, ..., 0]$, where 0 corresponds to the last cycle) with an exponential of the form: $\mathrm{NVR} = 1 + a10^{kb}$. The fit gives $a = 0.15 \pm 0.04$ and $b = 0.27^{+0.08}_{-0.05}$: the parameters are in good agreement

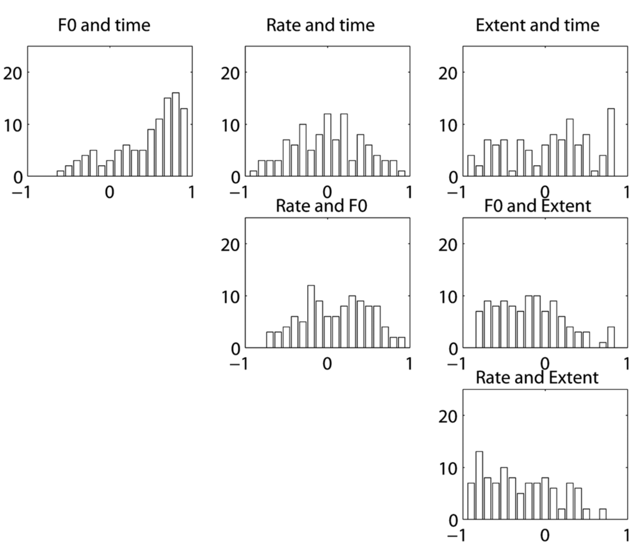

FIG. 5. Distribution of the correlation (in each tone) among vibrato parameters or vibrato parameters and time. The most striking features are the positive correlation between $F0$ and time (the singer increases the pitch while holding the note) and the negative one among rate and extent. A bit less significant is the dominance of negative correlation among $F0$ and VE.

TABLE III. Correlation among mean parameters of different recordings. The p-value of the smallest is reported in parenthesis.

	MVR	MVE	Year
MVE	−0.62		
Year	−0.57	0.70	
Length	−0.22	0.30	0.40
	$p < 2.4\%$	$p < 0.2\%$	$p < 2 \times 10^{-5}$

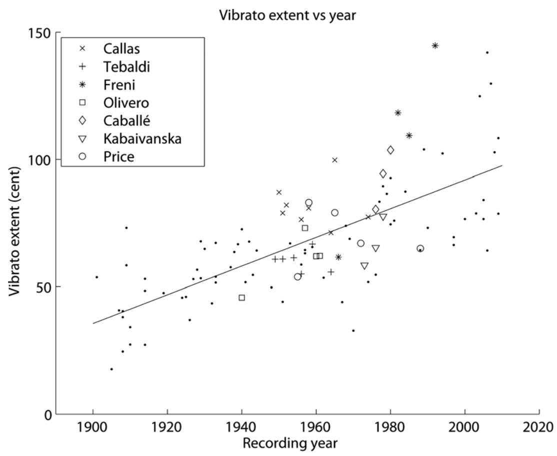

FIG. 7. Mean vibrato extent vs recording date. The line shown is a fit to all data.

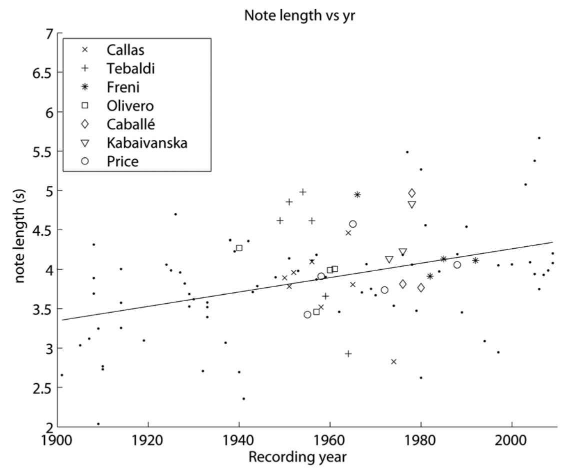

FIG. 9. Note length as function of recording date. A small increase (0.9 s/century) due to the evolving of performance practice can be seen.

with the ones found by Bretos and Sundberg but differs somewhat from the findings of Prame, who quotes $b = 0.59$.

The advantage in having data recorded over a large time span is that some historical trends in vibrato parameters can be observed. It has long been known[4,6,10] that at the beginning of the century faster vibrato rates were preferred: Figure 6 confirms that: one can easily see that the data, taken as a whole, indeed show a clear decrease of the measured vibrato rate as function of recording date: older performances at the beginning of the century seem to favor high vibrato rate, around 7 Hz, while most of the recent recordings show a vibrato rate around 5 Hz. One can also see that when multiple recordings of the same singer are analyzed, a clear decrease in vibrato rate due to aging can be seen in most cases. A linear fit to data gives a slope of 1.8 ± 0.3 Hz/century. The correlation coefficient is $r = -0.57$ (see Table III).

A similar plot (Fig. 7) shows that older recordings present a smaller vibrato extent (about a quarter of tone) with respect to modern ones: the fit gives a slope of 56.4 ± 0.3 cent/century. The correlation coefficient is 0.70.

Not surprisingly, one can the observe a clear negative correlation between vibrato rate and extent: see Fig. 8. If one uses vibrato amplitude VA, measured in Hz, one can try to fit data with an empirical law of the type $\mathrm{VA} = C(\mathrm{VR}^n)$, which gives $n = -1.6 \pm 0.7$. Forcing an integer n, the preferred value is $n = -2$.

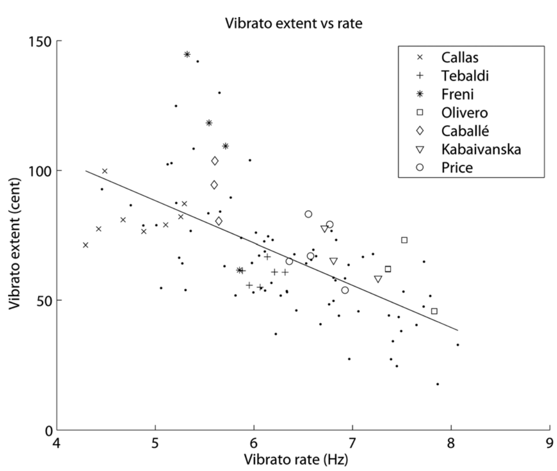

FIG. 8. Vibrato extent vs rate. A clear correlation (-0.62) can be seen. A linear plot works well for most data, except some low rate, high extent ones. The slope of the fit is -16.3 ± 0.1 cent/Hz.

One can also observe (Fig. 9) a positive trend also on note length, with an average increase of about 0.9 s/century: no fermata is marked on the note, and indeed older recordings tend to preserve its musical value. However, performance practice in the later years of the 20th century asks the singer to stop and hold the high tones as long as possible. The shortest note is by Farrar (2.1 s) while the longest one is by Cedolins (5.7 s).

Correlation of note length with vibrato extent, as found by Prame[5] can also be seen: the correlation coefficient is 0.30 with a significance of 0.1%.

V. DISCUSSION

The decrease of vibrato rate over the last century is a well known phenomenon whose causes are still uncertain. It is interesting to note that the rate at which change occurs is similar to the one observed in individual singers and due to aging.[11] It could be that during the training, singers try to imitate the vibrato rate and amplitude of their teachers, or of their older colleagues. It could also be due to the diffusion of recordings: voices with fast and narrow vibrato could sound "old" to students, who, in the quest of a more modern sound, could develop a slower, ample vibrato. It will be interesting to see what will happen next. In 1987 Sundberg wrote: "generally a vibrato rate of less that 5.5 undulations per second sounds unacceptably low";[1] however, most modern singers have a vibrato rate which is slower than 5.5 Hz and audiences do not complain. Will the trend continue, so that we will find acceptable even sounds which are now perceived as wobbling? Or there will be a discontinuity or a trend inversion which will bring singers again toward high rates?

VI. CONCLUSIONS

The work presented in this paper is an investigation of 105 recordings performed by 75 singers of the extreme high note in *Vissi d'arte* from *Tosca* by Giacomo Puccini. All artists were trained in the Western classical tradition and performed primary roles on important stages around the world.

The results obtained confirm the findings of previous works, but gives statistical significance or quantitative evidence to some results. In particular, the well know slowing down of vibrato rate observed from the dawn of voice recordings to our days can be observed with striking evidence: a constant decrease of 1.8 ± 0.3 Hz per century, without jumps or change of slope can be measured. Moreover, vibrato extent also shows an increase of 56.3 ± 0.1 cent/century of the same nature. Different behaviors could, in principle, be seen in different repertories (Baroque, Rossini) or in different singing styles. A correlation of vibrato rate and extent, supposed but never clearly demonstrated, can also be seen.

ACKNOWLEDGMENTS

I wish to thank Professor Johan Sundberg and Professor Steve Shore for their encouragement.

[1] J. Sundberg, *The Science of the Singing Voice* (Northern Illinois University Press, DeKalb, IL, 1987), Chap. 8, pp. 163–176.

[2] C. E. Seashore, *Psychology of Music* (McGraw Hill, New York, 1938), Chap. 4, pp. 33–52.

[3] A. Keidar, I. Titze, and C. Timberlake, "Vibrato characteristics of tenors singing high c's," in *Transcripts of the 13th symposium Care of the Professional Voice, Part I*, edited by V. Laurence (The Voice Foundation, New York, 1984), pp. 105–110.

[4] E. Prame, "Measurement of the vibrato rate of ten singers" J. Acoust. Soc. Am. **96**, 1979–1984 (1994).

[5] E. Prame, "Vibrato extent and intonation in professional western lyric singing," J. Acoust. Soc. Am. **102**, 616–621 (1997).

[6] H. Rothmanm, J. Diza, and K. Vincent, "Comparing historical and contemporary opera singers with historical and contemporary jewish cantors," J. Voice **14**, 205–214 (2000).

[7] Medea Mei Figner is best known as the first Lisa in Tchaikovsky's *Queen of Spades*.

[8] The Mathworks, Inc., "MATLAB®, the language of technical computing," version 6.5 (2002).

[9] J. Bretos and J. Sundberg, "Vibrato extent and intonation in professional western lyric singing" J. Voice **17**, 343–352 (2003).

[10] A. Reinders, "The history of vibrato in the singing voice," in *Vibrato*, edited by P. Dejonckere, M. Hirano, and J. Sundberg (Singular Publishing Group, San Diego, 1995), Chap. 8, pp. 141–144.

[11] J. Sundberg, M. N. Thörnvik, and A. M. Söderström, "Age and voice quality in professional singers," Logoped. Phoniatr. Vocol. **23**, 169–176 (1998).