A method for extracting vibrato parameters applied to violin performance

Julieta Gleiser, Anders Friberg and Svante Granqvist

Abstract

A method is presented which semi-automatically extracts the fundamental frequency and displays as continuous signals vibrato rate, vibrato extent and sound level. The method is tested on specially made recordings of violin music with piano accompaniment, using a small microphone mounted directly on the violin. The fundamental frequency was successfully extracted by means of a waveform correlation program. Likewise, vibrato rate and extent were extracted separately for each tone from the fundamental frequency signal after elimination of its DC component. The results seem promising, offering the opportunity of visual examination and measurement of changes in vibrato characteristics during performances of entire pieces of music.

Introduction

Frequency vibrato is an essential aspect of music performance, typically occurring in singing and in bowed instruments. It consists basically of a modulation of the fundamental frequency F0. Both the modulation amplitude and frequency, generally referred to as extent and rate, have been found to be perceptually salient parameters. Also the sound level is mostly modulated because of the resonance properties of the instrument. In current synthesizer technology complete reproductions or emulations of real instrument sounds are often used. A numerical analysis of vibrato characteristics can be expected to yield important contributions to synthesis, particularly in the case of bowed instrument sounds.

Generally these parameters have been measured manually by examining the F0 modulation waveform. The results have shown that vibrato rates differ between different performers (Seashore, 1936). Recently, De Poli and co-workers (1998). determined means of these parameters by matching a sinusoid to the modulation waveform. They found that the rate, averaged over the tones in a piece, differed significantly depending on the player's expressive intention. In addition, their data revealed that the rate varied considerably between tones. Prame (1994, 1997) measured the vibrato in singing manually from spectrograms in 10 commercial recordings of Schubert's *Ave Maria*. He observed that the vibrato rate typically was increased toward tone endings. These investigations suggest that the vibrato varies both between and within tones in a fashion that may be related to musical expression, thus motivating a more detailed analysis. For such an analysis, a manual method of measurement is highly time consuming.

This investigation presents a semi-automatic method for analysis of three aspects of the vibrato: rate, extent and sound level fluctuation. The method has been tested on music played on the violin, the overall aim being to relate the variation in these parameters to musical expression. In future work, we intend to analyse how vibrato parameters relate to musical structure and expression.

Recordings

To increase accuracy of measurement and number of examples and to include the analysis of sound level, we decided to analyse laboratory recordings rather than commercial recordings. A special recording technique, allowing the isolation of the violin tones from the accompaniment, was used. It also made it possible to manipulate the signal by means of special developed software.

Six pieces from the Western classical music repertoire were selected by Lars Frydén (Table 1). The pieces differed in the character in such a way that vibrato differences could be expected. Also, since vibrato in bowed instruments typically occurs on long notes, the pieces were in a slow or moderate tempo.

Three violin players performed these pieces with piano accompaniment. Violinist 1 is a toplevel student and had recently graduated from the Edsberg Institute of Music. Violinists 2 and 3 are Concert Masters in professional symphonic orchestras in Stockholm. The players selected professional accompanists according to their preferences. The recording yielded a total of 18 examples (3 players x 6 pieces).

Table 1. Pieces selected for vibrato analysis, all arranged for violin and piano.

Composer	Piece
F. Schubert	*Ave Maria* in G major, op. 52 No. 4 (Edition Peters),
F. Kreisler	*Melodie* (Charles Foley, New York)
W. A. Mozart	*Adagio* from Violin Concerto No. 3, KV 216 (Edition Peters)
J. Haydn	*Andante Cantabile* from String Quartet in F major, Op. 3, No. 5
Padre Martini - Kreisler	*Andantino* (Carl Fischer, New York)
J. S. Bach	*Air* from Orchestral Suite No. 3 in D major, BWV 1068

The recordings were made in a room used for exhibitions and seminars at TMH, 9 x 9 x 3 m, approximately. The accompaniment was played on a Steinway C (224 cm long) grand piano. The microphone used for the violin was a Brüel & Kjær Cardioid compact Mic, Model 4021. It was placed in a holder (VH 4000) attached to the strings behind the bridge of the violin. A Brüel & Kjær 4003 Studio microphone was used for the piano; it was placed in the 10-cm high-lid opening. A cover was placed over the piano to reduce leakage into the violin microphone. The microphones were connected to a Symetrix SX202 Microphone Preamplifier and then to each channel of a Sony TCD-D10 PRO Digital Audio Tape recorder. The recordings were digitally transferred to a PC as ".wav" files, using a sample rate of 44.1 kHz.

The musicians were given the music scores in advance and at the recording session they were again given the opportunity to rehearse before the recording was made. They were asked to play as during a concert but were not given any specific instructions on how to perform. In particular, no reference was made to vibrato. Each piece was repeated until the musicians were satisfied with the performance. This last version was used for the subsequent analysis.

Extraction of the vibrato parameters

The vibrato parameters were obtained mainly by using programs included in the Soundswell Signal Workstation and the Spruce package from AB Nyvalla DSP. The sound files and the resulting analysis files were all in the so-called SMP format. The procedure for obtaining vibrato extent, vibrato rate and sound level was the following.

1. Each file was split into minor sections (4 to 8 parts). This was done by placing labels in the program *Swell* and cutting the file into smaller parts with the custom made program *Cut*.

2. By running the *Corr* program, a correlogram was obtained for each file. A correlogram (Granqvist, 1998) shows the Pearson correlation coefficient between the signal portions in two time windows, one fixed and one moving. The vertical axis represents the inverse of the time separation between the windows. The grey scale shows the correlation after optional application of compression function, characterised by a gamma parameter. The horizontal axis represents the time co-ordinate for the beginning of the fixed window. A correlogram thus shows in terms of dark curves possible candidates for the fundamental frequency versus time. With a specially developed tracing feature the curve corresponding to the F0 was manually selected and an F0 extraction was made (Figure 1). The settings selected for this purpose included a window length of 10.0 ms, a cell size of 1 pixel, the gamma parameter set to 4.0 and the threshold to 90.3 [dB]. The resulting F0, downsampled 10 times, was saved to an SMP-file.

3. The short SMP files, each containing F0 extractions for a section of the piece, were glued together to one single SMP using the *Glue* program. The F0 data were converted from Hertz to cent relatively A440 by means of the *Extract* program. Figure 2 shows an example of the resulting file showing violinist's F0 expressed in cent.

4. The sections containing vibrato were manually selected and isolated from the rest of the signal. The mean value of F0 in each section was computed using the *Hist* program. This
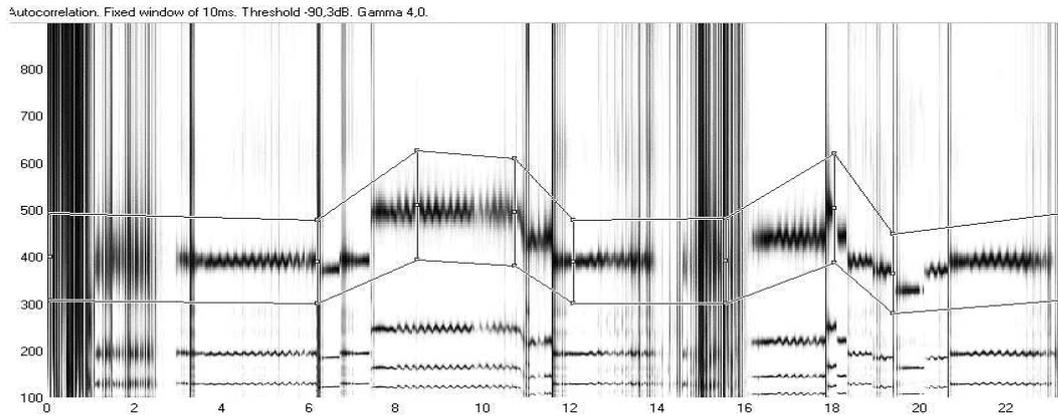
Figure 1. Example of a correlogram showing frequency versus time (see text), with the manual trace marking used for the F0 extraction. This graph corresponds to the first two phrases of F. Schubert's "Ave Maria".
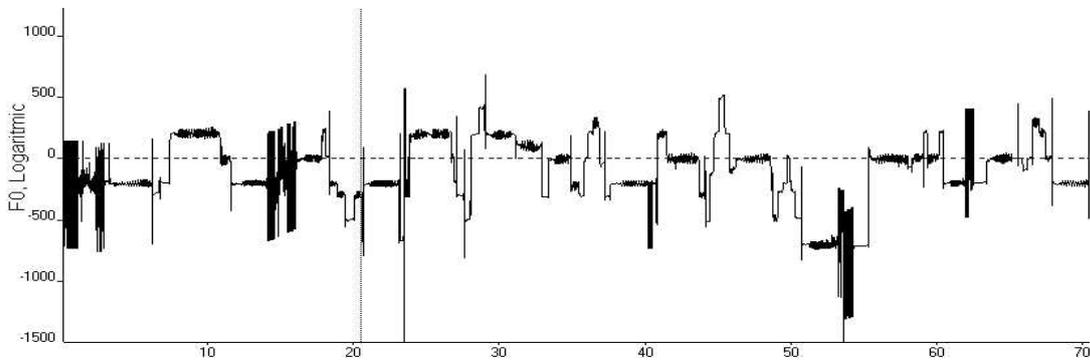

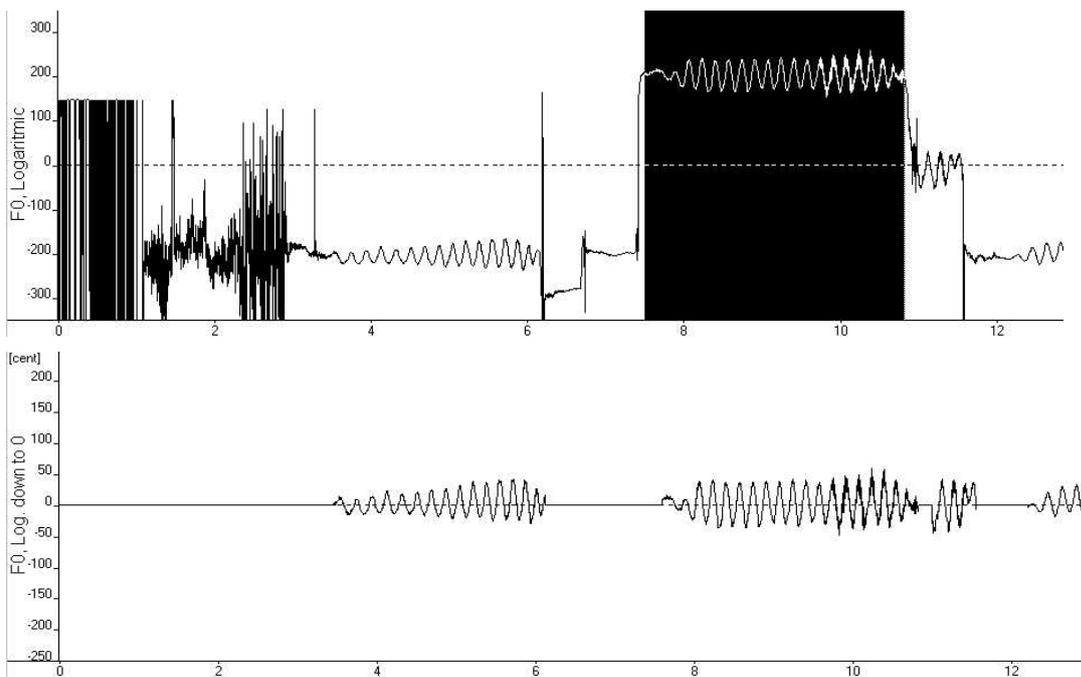*Figure 2. The resulting F0 extraction in cent. (From Schubert's "Ave Maria").*

Figure 3. Isolation and selection of sections containing vibrato. The upper and lower graphs show the F0 signal and the resulting isolated vibrato, respectively.
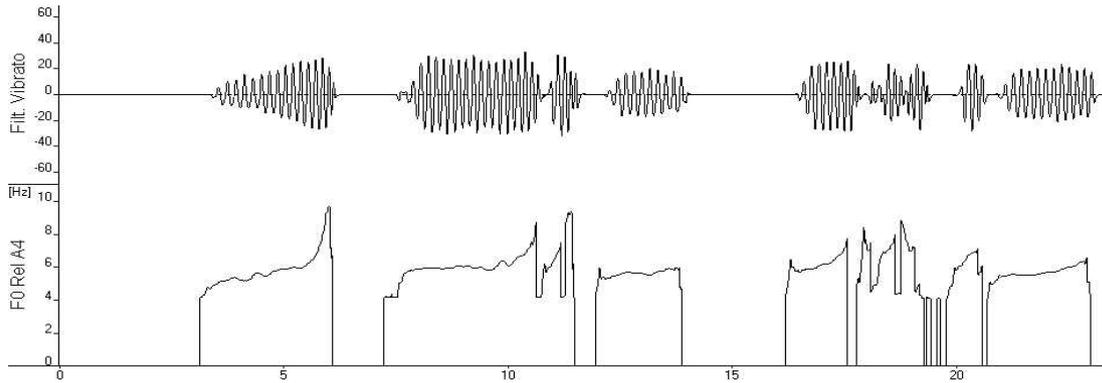
Figure 4. Vibrato and vibrato rate (upper and lower graph) for the first two phrases of F. Schubert's"Ave Maria" as performed by violinist 2.

so that for each vibrato section a value of zero corresponded to the average F0. The signal for non-vibrato sections was set to zero. The upper and lower panels in Figure 3 show the original F0 signal and the resulting isolated vibrato signal, respectively. Sometimes, a low sound level in the beginning or end of a tone prevented F0 tracing. An example can be seen in the beginning of the first tone in Figure 3. Such parts of the tone were excluded. Interestingly, the sound level at the end of the last note of the first phrase in *Ave Maria* was sometimes so low that the tone ending could not even be detected aurally in the recording, even though the microphone was about 5 cm from the string.

5. The isolated vibrato signal was then high and low pass filtered, again using the *Extract* program. Each of these filters comprised two second order Butterworth filters, applied both forward and backward, so as to obtain a zero phase shift. A cutoff frequency of 20 Hz was set for the lowpass filter, and a cutoff frequency of 4 Hz for the highpass filter.

6. The *Corr* program was then applied again, this time to extract the vibrato rate (Figure 4). This had the advantage that a value for

the vibrato rate was obtained for each sample. This is different from previous investigations where typically only the peaks and/or troughs of the waveform were detected. The settings where the same as those selected for the first F0 extraction, except for the frequency scale and the window length, the latter being set such that it corresponded exactly to one vibrato period. The program picked the starting point of the first, fixed window as time co-ordinate for the maximum correlation. This means that each vibrato rate point in Figure 4 refers to the following two vibrato periods.

7. The vibrato extent in cent was computed with *Extract* using the level option (Figure 5). An RMS value with a full wave rectification and a lowpass filtering of 6 Hz with zero phase shift was selected.

8. Finally, the sound level was extracted (Figure 6). An RMS value with a full wave rectification and a lowpass filtering of 200 Hz with zero phase was chosen.

Examples of the complete set of graphs are shown in Figures 7, 8 and 9 for the complete *"Ave Maria",* as performed by violinists 1, 2, and 3, respectively.
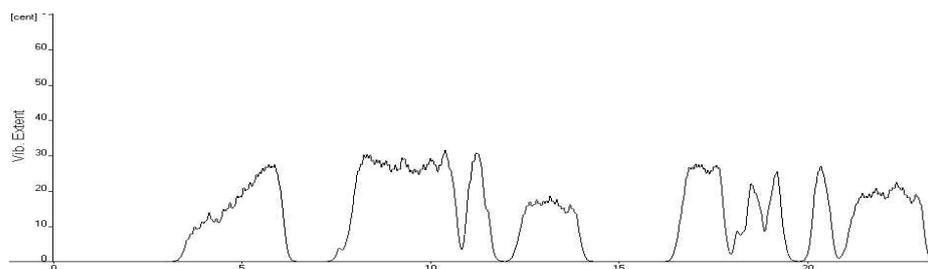

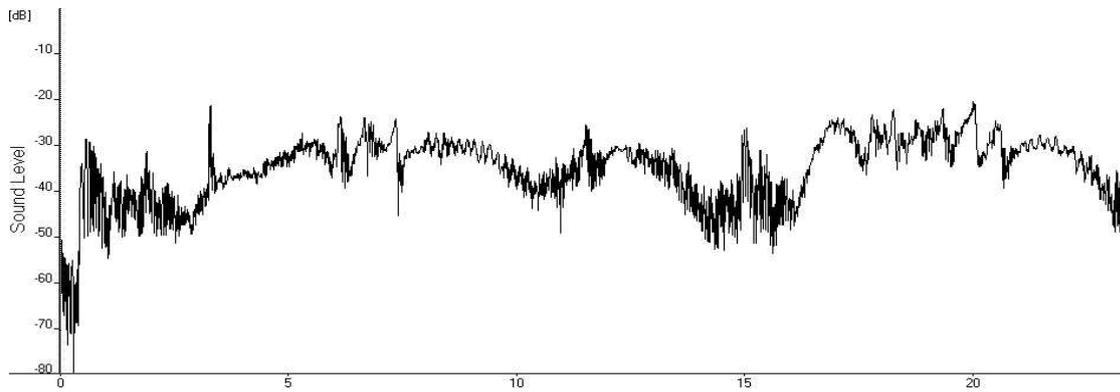*Figure 5. Resulting vibrato extent. Same example as Figure 4.*

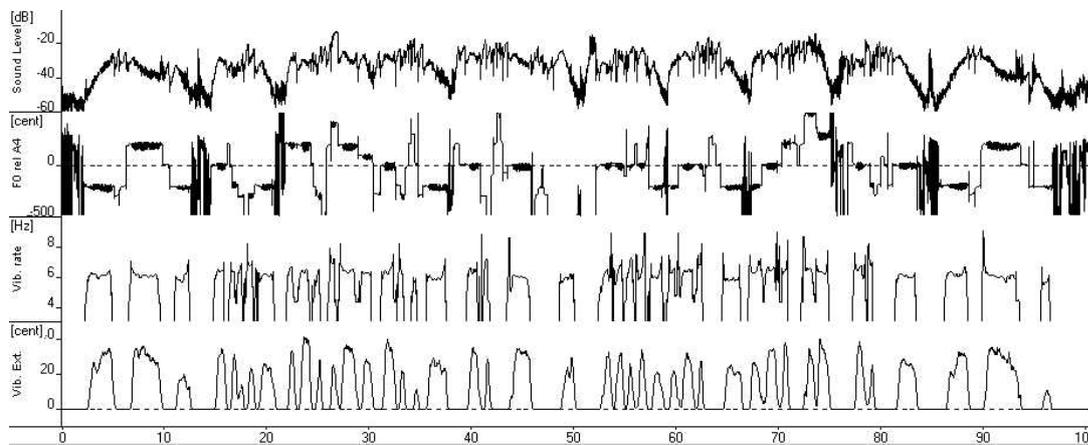*Figure 6. Sound level of the example shown in Figure 4.*

Figure 7. Parameter graphs for the complete version of "Ave Maria", as performed by the violinist 1. The four graphs show, from top to bottom, as a function of time in seconds: sound level, F0, vibrato rate, and vibrato extent given in cent as the positive and negative peak departures from the mean F0.
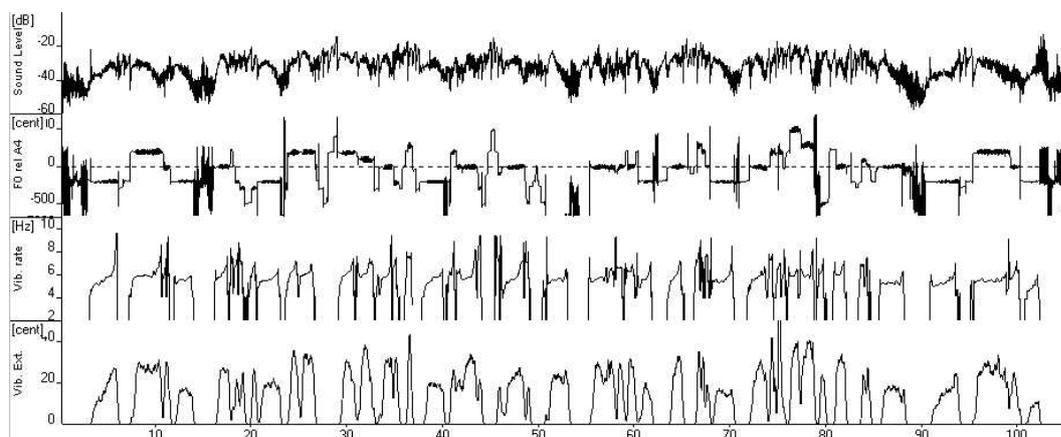

Figure 8. Same example as shown in Figure 7, performed by violinist 2.
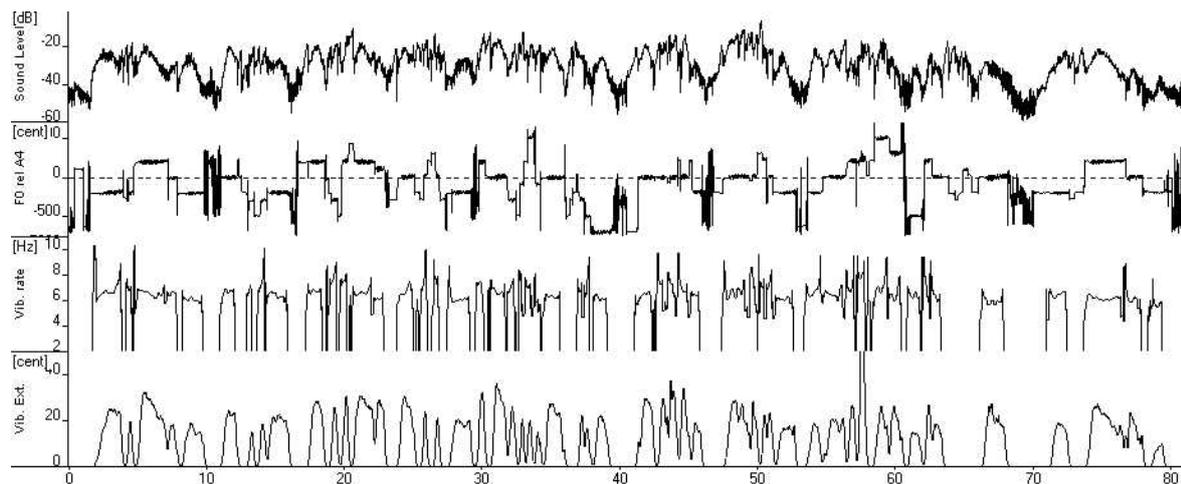
Figure 9. Same example as shown in Figure 7, performed by violinist 3.

Although we intend to make a detailed analysis in music performance in the future, some examples of vibrato characteristics revealed by the present method can be mentioned. Many tones end with an increasing vibrato rate, especially in the case of violinist 2 (Figure 8). This is in agreement with that previously observed for singers (Prame, 1994). The vibrato extent is often coupled to the sound level, so that the extent increases with increasing sound level, as in the first tone in Figure 8. In some cases, a correlation seems to exist also between sound level and the start of the vibrato, such that the vibrato onset is delayed until the sound level has reached a peak, see second and third vibrato tones in Figure 7. However, these vibrato para-meters do not appear to be generally coupled; the players seem capable of varying them independently, according to the demands of musical expression.

In summary, the results obtained so far seem promising. In particular, the method used for F0 and vibrato rate extraction seems quite reliable. The method displays vibrato characteristics as function of time and thus appears capable of revealing how such characteristics are varied in music performance.

Acknowledgements

References

De Poli G, Rodà A & Vidolin A (1998). Note-by-note analysis of the influence of expressive intentions and musical structure in violin performance. *Journal of New Music Research* 27/3.

Granqvist S (1998). Correlogram, a visual display of periodicity. *TMH-QPSR, KTH* 4/1998: 13-22.

Prame E (1994). Measurements of the vibrato rate of ten singers. *J Acoust Soc Am* 96/4: 1979-1984.

Prame E (1997). Vibrato extent and intonation in professional Western lyric singing. *J Acoust Soc Am* 102/1: 616-621.

Seashore CE (1936). Psychology of the vibrato in voice and instrument. In: *University of Iowa Studies in the Psychology of Music, Vol. III* (Univ. of Iowa, Ames).