ACOUSTICS 2012

The impact of vibrato usage on the perception of pitch in early music compared to grand opera

H. Daffern, J. Brereton and D. M. Howard

University of York, Audio Lab, Electronics department, Heslington, YO10 5DD York, UK
helena.daffern@york.ac.uk

Previous studies on the pitch of long-duration vibrato tones have typically used synthesised modulator tones to assess the perceived pitch in relation to its arithmetic or geometric mean fundamental frequency. In this study a listening test was conducted with expert listener subjects matching recorded vibrato tones sung by professional singers using a method of adjustment and free response paradigm as employed in previous research. Example tones selected from recordings by 16 singers were used in the test, 8 of whom were employed at the Royal Opera House, Covent Garden, and the remaining singers specialised in Early Music Performance. Previous research into vibrato usage by these singers shows a noticeable difference in the use of vibrato between these performance groups, particularly in terms of extent throughout long tones. The impact of these differences in vibrato is assessed in terms of the perception of fundamental frequency in long tones as performed by the two groups of singer.

1 Introduction

Vibrato is known to be an important factor in the measure of vocal quality, particularly in Western classical singing. A number of studies in vibrato production of professional opera singers have produced common standards of vibrato rate and extent associated with good singing. A vibrato rate of 5 – 8 Hz and peak to peak extent of up to a tone are generally considered to be indicative of good vocal quality [1, 2]. The regularity of the vibrato is also significant: based on combined analysis of vibrato tones with listening tests, Diaz et al. found 'a direct relationship between the periodicity of the vibrato wave and its perceived quality'[3].

Vibrato has also been found to have characteristic treatments in different genres of classical singing. Daffern conducted a study comparing professional singers of grand opera compared to professional singers specialising in early music and found that there were noticeable differences in vibrato production between the two groups of singer [4]. In addition to slightly larger average vibrato extent employed by the opera singers compared to the early music singers, there was a noticeable difference in the production of vibrato throughout long tones. The trend was for opera singers to begin vibrato production with the onset of a tone, whilst the early music singers would often produce a large proportion of the note as an essentially 'straight' tone before adding vibrato as a stylistic affect.

Vibrato studies tend to focus either on the analysis of vibrato produced by real singers, to assess acoustic factors such as rate, extent, regularity and physical factors such as the production of vibrato by singers [5,6,7,8], or on the perception of vibrato by listeners[9,10,11]. Howes et al. conducted studies which combined vibrato analysis of recorded sung tones with listening tests, focussing on the listener's aesthetic reaction rather than perception of the analysed characteristics of vibrato [2]. Perception studies in vibrato rely heavily on listening tests and often use synthesised tones to allow the many variable features of vibrato to be controlled.

One important factor of vibrato is the perception of pitch associated with a vibrato tone. Research using synthesised tones suggests that listeners perceive the mean fundamental frequency as the pitch of a long vibrato tone. However, Sundberg [12] found evidence that the linear average of the fundamental frequency contour was significant, whilst Shonle & Horan [13] found the mean of the extreme fundamental frequencies in a vibrato tone to be more consistent with perceived pitch. A similar study that used short-duration vibrato tones of 0.5 to 2 vibrato cycles found the fundamental to 'correspond to a weighted time average of the F0 pattern' and not the mean of the frequency of the peaks, they also found that the 'pitch depends most on the ends of tones'[14]. Van Besouw et al. used a different protocol to d'Alessandro, in which the subject heard a modulated tone of one second before and after the tone being tuned by the listener and found that 'participants were more influenced by the initial phase of the second modulated tone than the end phase of the first' [9]. Gockel et al.[15] presented a model based on the rate of frequency change on the overall pitch, based on the theory of perceived pitch of vibrato tones corresponding to time weighted aspects of the F0 across a vibrato tone.

Recorded rather than synthesised tones have been used by Brown et al. [16]: tones of approximately one second played on a viola were used to perform a pitch matching listening test using experienced musician participants. The results were found to be consistent with the synthesised experiments using long tones in that participants matched to the mean fundamental frequency. Based on the research of Brown et al., Yoo et al. conducted a study using recorded violin tones and found that 'one-second length of vibrato tone is more than enough for subjects to be able to resolve the tone to its mean pitch' [17].

The pitch perception of recorded sung tones rather than synthesised vowels has not yet been considered in detail, and in particular the effect of different treatments of vibrato throughout long tones which is commonplace in modern classical performance conventions has so far been overlooked.

2 Experiment

2.1 Listening test

Participants were required to listen to a set of vibrato tones and tune a second, synthesized straight tone to the perceived pitch of the first. A method of adjustment was used based on the experimental protocol by van Besouw [9] but in this case using recorded tones rather than synthesised stimuli. A Contour Designs ShuttleXpress (2003) allowed listeners to adjust the step size and gradually tune an adjustment tone to the pitch of the vibrato tone by rotating the wheel in the centre of the shuttle (Figure 1). Whilst in van Besouw's study the modulated (vibrato) tone was played before and after the adjustment tone. In this experiment the modulated tone was played only once, before the straight tone. This sequence (vibrato – straight tone) could be played multiple times until the participant was satisfied with the match. As in van Besouw's experiment the initial frequency of the adjustment tones was 15 – 20% above or below the centre frequency of the vibrato tone (in this case C5), which is the written pitch of the recorded tone.

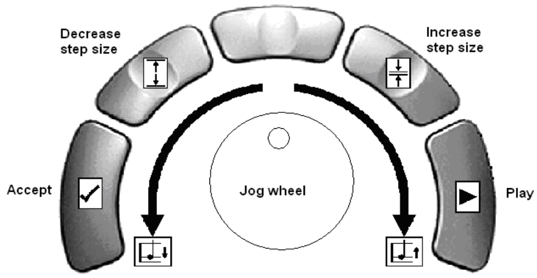

Figure 1: Diagram of the shuttle device used to play, tune and accept adjustment tones Adapted from van Besouw [9]

The tones were presented in a random order to the participants over Beyer Dynamic DT100 Headphones. Each tone was presented twice, with the adjustment tone above and below the mean F0 of the vibrato tone. Participants were able to adjust the volume and were encouraged to take breaks when needed. The number of repetitions before a match was made and the time taken for each match was also recorded. Any difficulties in pitch matching specific tones were reported to the experimenter during the test. The experiment was conducted in the Listening Room of the Audio Lab at the University of York.

2.2 Vibrato tones

The first note of Handel's 'Ombra Mai Fu' was selected from the recordings made by Daffern [4] which include the first section of the aria sung unaccompanied by 16 professional singers, eight specialising in opera (O) and eight specialising in early music (EM). Four vibrato tones were selected from the data, two from each group of singers which were idiomatic of the vibrato characteristics for each genre based on the results of the study conducted by Daffern (2008). S1 and S2 represent EM and S3 and S4 the O group.

From each selected tone four example tones were created: the note in its entirety as produced by the singer (Sx-ALL); a one second sample taken from the beginning of the tone after the onset (Sx-BEG); a one second sample taken from the middle of the tone (Sx-MID); a one second sample taken from the end of the tone before the offset (Sx-END). Figure 2 presents F0 contours of the tones selected, illustrating the vibrato characteristics for each tone.

2.3 Participants

Five participants took part in the study, all enrolled in full time music degree courses at the University of York. Four of the five participants were female and all were first study singers aged between 19 and 24. Subject 4 was also heavily involved in conducting choirs and was the only postgraduate student. All participants reported to be in good health on the day of the study and with no history of any hearing impairment.

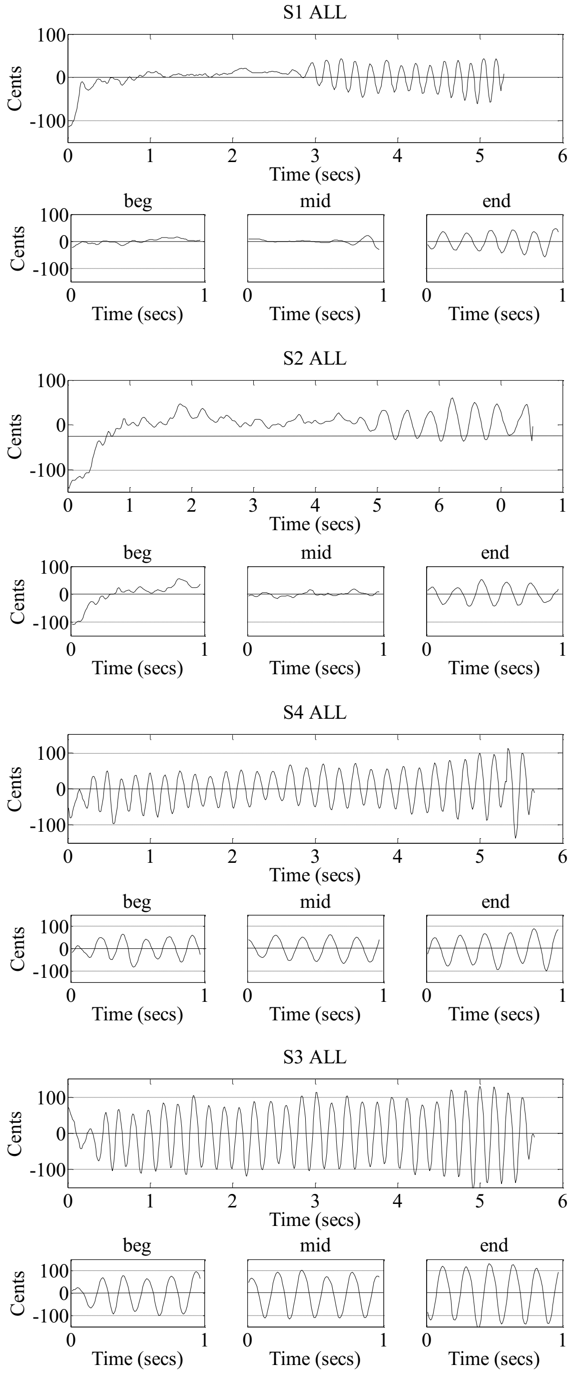

Figure 2: Vibrato plots showing the F0 contours of the selected tones for S1& S2 (EM) and S3 & S4 (O)

3 Analysis

F0 data was extracted from the audio files of the vibrato tones using the voice analysis software PRAAT (http://www.fon.hum.uva.nl/praat/) and was then transferred to MATLAB (http://www.mathworks.co.uk/) for vibrato investigation and cross-examination with the participant's results from the listening test. As current research suggests that the mean F0 is the most significant influence on the pitch matching of vibrato tones of over one second, the mean F0 was calculated for all the test tones (Mv) and used as the basis for analysing the results of the listening test (F0 of the entire tone was calculated irrespective of the vibrato content of the tone).

The mean F0 for each tone is presented in Table 1.

Table 1: Showing the mean F0 (Mv) of the vibrato tones

	Mv of Tone			
Singer	ALL	BEG	MID	END
1	520Hz	520Hz	519Hz	524Hz
2	525Hz	522Hz	527Hz	527Hz
3 (O)	520Hz	517Hz	524Hz	520Hz
4 (O)	517Hz	513Hz	520Hz	521Hz

The data from the participants (P) was considered based on their accepted pitch match (Pm) compared to Mv and the difference between Pm to Mv was calculated in cents (Md).

4 Results

All subjects make matches close to Mv: the mean Md across all subjects is 13.8 cents, suggesting that subjects were matching to Mv as expected. However, trends of accuracy in pitch matching to the mean are apparent when the results are considered between conditions and participants. Table 2 shows the mean Md for all matches for each participant alongside the averaged results for the O group and EM group. P5 produced matches furthest from the Mv (Mean Md = 23.8 cents), whilst P2 made the closest matches overall (Mean Md = 7.7cents).

Table 2: The mean Md in cents for all participants (P) with O group and EM group tones also calculated separately

P	O	EM	All tones
	Mean	Mean	Mean
1	12.6	8.3	10.4
2	7.4	8	7.7
3	14.9	13.1	14
4	18.4	9.6	9.5
5	23.7	22.8	23.8
All	15.4	12.3	13.8

4.1 Tone length

The length of the tone does not appear to effect the accuracy of pitch matching to Mv, the mean Md of all subjects for long tones being 13 cents compared to 14 cents for the short tones (it is important to note that only 4 long tones were used in comparison to the 12 short tones which will have an impact on the averaging of the data). This implies that an averaging effect over the whole length of a tone is taking place and that once tones are at least one second long there is a high level of accuracy in matching to the averaged F0. The contrast in the treatment of vibrato and effect on F0 between the long and short tones makes this a surprising finding: the mean F0 values for short tones in Table 1 show an overall upward shift in F0 for the long tones for all the singers. If listeners are influenced by the moving mean F0 of the tone, the expectation would be for the participants to match either above or below the mean consistently depending on the influential part of the tone. However, participants in this study matched long tones no more consistently either above or below the Mv.

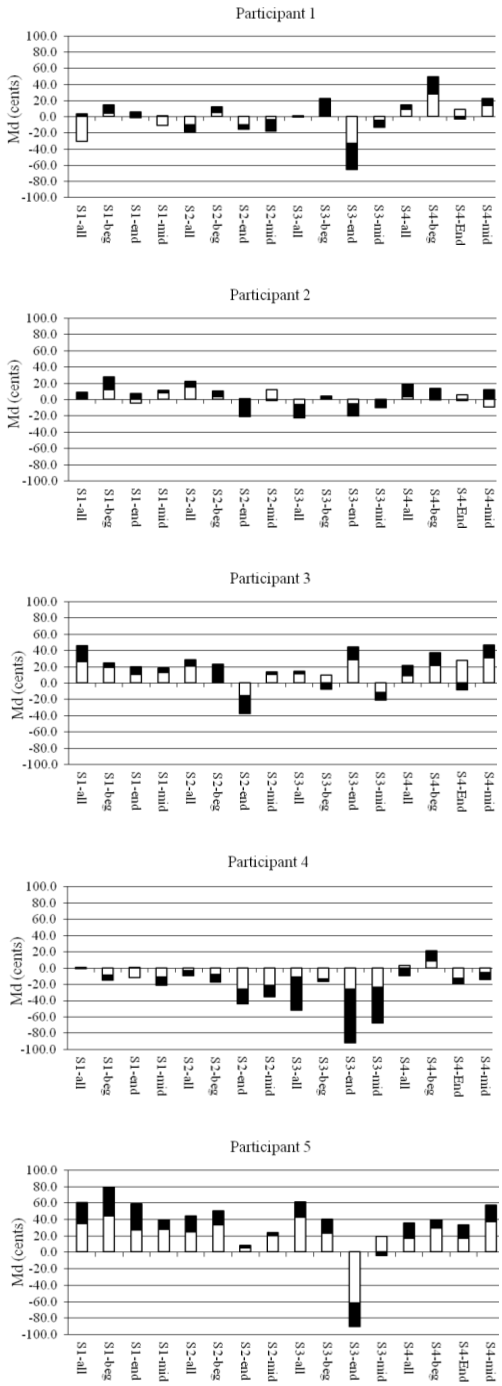

Figure 3: The Md for each tone matched by all Ps when the adjustment tone is from above (white) and below (black)

Figure 3 shows the Md for each participant matching with an adjustment tone from above and below. The accuracy of Pm to Mv does not appear to be correlated with the initial direction of the adjustment tone. P2 makes closer matches when the adjustment tone starts from above, however, the high accuracy of Md to Mv for this subject make this finding less noteworthy. The adjustment tone seems to have no influence on the direction of Md, in that, for most tones the Md is matched in the same direction for the same tone when adjusted from above and below.

S3-END produces the largest Md from each participant (Mean Md for S3-END = 27.38 cents). Consistent with results for the others tones P4 and P5 make the matches furthest from Mv at 61.9 and 66 cents respectively. This is the only tone with a vibrato extent that consistently exceeds a 100 cent deviation from the mean F0 (Figure 2). It seems that, although the vibrato is highly regular in this tone compared to other tones in terms of rate and extent, the large deviation of F0 is causing higher Md values. Four of the participants matched the tone below Mv which suggests that they might be matching to the beginning of the tone as the onset of the vibrato begins at the lower extreme of a cycle and ends at the upper extreme. The possible importance of the direction of match deviation to Mv is considered later in this paper.

4.2 Agreement of Md direction

Figure 4 represents the agreement of direction of deviation from Mv between subjects, showing the percentage of matches made higher and lower than Mv. There is a general trend for subjects to match more tones higher than Mv. Cross-examination of the patterns of individual participants in Figure 3 shows that P4 is the only participant not to follow this pattern. Whilst P1, P2 and P3 vary matches both higher and lower than Mv, P4 makes all but 2 matches well below Mv and P5 makes all but 2 matches much above Mv. The contrast in direction of deviation for these participants suggests that matching is listener specific rather than tone specific. However, the direction of deviation from Mv is often consistent between P2, P3 and P5, particularly for the EM tones. P4, due to the tendency of matching well below Mv provides the most discontinuity for a trend between participants.

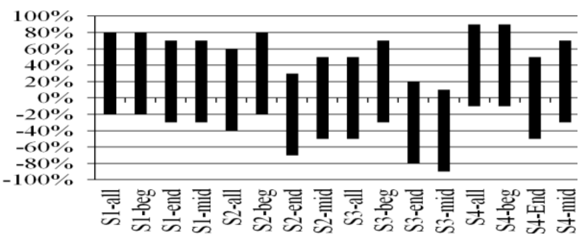

Figure 4: showing the % agreement between participants in matching higher (positive) or lower (negative) than Mv

Examination of the vibrato characteristics of the tones with most agreement when matching above or below Mv provides inconclusive evidence of being influential to the direction of the match, although some expected patterns do emerge. For example, S1-BEG was matched above Mv by all participants other than P1. Considering the nature of mean F0 throughout this tone, increasing considerably but with no periodic vibrato, a match higher than Mv is expected based on the suggestion that listeners match to the ends of tones [14].

Other emerging patterns are less consistent with expectation: S4-BEG, S1-END and S4-ALL were matched higher than Mv by four participants, with the anomaly provided once again by P4. These tones all begin with F0 values below Mv but with differing onset phases (i.e. direction of vibrato). S4-BEG and S4-All end close to but below Mv, whilst S1-END ends above Mv (although just beyond the peak of a cycle so that F0 is beginning to descend). Four participants were also in agreement matching higher than Mv for S1-BEG, however the lack of vibrato, alongside close proximity of the entire tone to F0 makes comparison ineffective.

S1-END also has a similar vibrato pattern to S3-END (although S3-END has a much larger extent) beginning below and ending close to the maximum extent in both case: however, in contrast to S1-END four participants matched this tone lower than Mv. The end phase does differ slightly between these tones as S3-END finishes toward the top of a cycle rather than at the start of a descent, however this does suggest that listeners are not matching to the start of the tone as previously suggested.

The two tones matched consistently below Mv by four participants (all except P4) were S3-MID and S2-END. Both these tones begin and end above Mv, although very close to Mv for S2-END, with an initial phase that is consistent with the final ascending portion of a vibrato cycle for the tone. Based on the tones examined, the direction of the matches to Mv appear to oppose the direction of the beginning or end of the tone, in that when the tone has a beginning and end higher than Mv the match is lower and vice versa. The lack of agreement between participants on direction of Md for S4-END considering the similar patterns of vibrato in terms of onset and offset phase and the similar mean F0s for this tone and S3-MID reduce the importance of this result.

It is difficult to assert based on this data set whether there is a correlation between the characteristics of specific vibrato tones and the direction of Md from Mv: although there was agreement for many of the tones, the absence of more data makes it difficult to draw noteworthy conclusions, especially considering the variation in the onset and offset phases of the tones.

4.3 Impact of genre on Md

The genre of the vibrato tone (O or EM) does have some impact on the results of some participants, however, the effect is not consistent. The mean Md across all participants when matching tones from the O group is slightly higher than for the EM group, at 15.4 and 12.4 cents respectively. However, consideration of the individual participant data reveals that this is not representative of individuals. P2 and P5 show very little discrimination between genres with less than 1 cent difference in matching and whilst P1 and P4 make closer matches for the EM group vibrato tones (8.3 and 9.3 cents respectively) than the O group (12.6 and 18.4 cents respectively), P3 makes much closer matches for the O group (12.6 cents) than EM group (22 cents). This suggests that the characteristics of vibrato from the two genres impact on the pitch matching of the mean F0 but is dependent on the listener rather than the characteristics themselves.

As Figure 2 shows, the characteristics of the whole vibrato tones is such that once edited into short tones the features alter and often become indistinguishable across genre. For example, the END tones for the EM group (S1 and S2) are very similar to the BEG tones of the O group as the 'straight' portions from the ALL tones in the EM group have been removed. As the introduction of vibrato well after the onset in long tones is idiomatic of early music singing [4], the matching of the short EM tones, which do not reflect this characteristic, that is S1-END and S2-END, would be expected to correspond more closely to the matches in the O group than other matches in the EM group. The Md for S2-END P4 is the largest for the EM matches for this participant and is in line with the matches to the O tones as expected. However, the Md for S1-END is consistent with the other EM matches for this participant. P1, who also produced lower mean Mds overall for the EM group also makes matches in line with other EM matches for S1-END and S2-END which would not be expected. These inconsistencies in terms of the Md relating to the vibrato characteristics of the tone suggest that a factor other than the vibrato behaviour throughout a tone is contributing to the pitch matching. The lower extent of the END vibrato tone for the EM group could partly explain this result.

4.4 Participant feedback

Participants gave verbal feedback sometimes during the test to discuss specific tones as well as at the end of the test. All 5 participants commented on the difficulty of matching the EM long tones: ‘long notes with a straight beginning were hardest’ (P4). Four of the five participants found tones with ‘a big vibrato’ (P1) difficult, P5 commenting on the vibrato entering immediately with S10-END finding it ‘so wobbly I don’t know what she’s going for’ (P5).

When identifying specific notes within the test as being difficult to match, there was rarely duplication between subjects. Tones ‘flagged up’ by more than one participant include S1-ALL, S2-BEG, S3-END and S4-ALL. S2-BEG was the only tone to be identified both times it was played. The obvious difficulty subjects had with this tone is unsurprising considering the dramatic shift in mean F0 throughout the tone combined with the lack of vibrato. However, the Md results for this tone for all participants are consistent with other results. This lack of correlation between the Md results and the difficulty observed by the participants is true of all identified tones apart from S3-END which was discussed above. This suggests that whilst participants may feel distracted or have to concentrate harder, the pitch they associate with the tone remains the mean F0 averaged over the whole tone.

5 Conclusions

Overall, the close matching to Mv by participants suggests that, in line with current thinking, listeners do match to the mean of the vibrato tone. The inter-subject and cross-subject patterns emerging from examination of the direction of Md suggest that certain factors within a tone influence the match to be biased either above or below Mv. The consistent matching of P4 and P5 below and above the mean respectively, alongside the larger Mv values for these subjects suggests that Mv direction is dependent upon the listener. However, the commonalities across all participants excluding P4 suggest that it may be possible to formulate some generalisations. The large variations in the vibrato characteristics of the tones combined with the small number of participants make it difficult to draw conclusions as to which vibrato features are affecting at this stage. The lower mean values of Md observed for EM tones combined with the drastically larger Md results for S3-END suggests that a larger vibrato extent (consistent with O group) influences matches to deviate further from the mean F0. However, although not obviously reflected in the results, the agreement of participants that the ‘straight’ tones were hardest to match suggests that vibrato can be helpful when assigning a pitch to a tone, although there is likely to be an optimum extent beyond which it becomes detrimental to pitch matching to the mean.

Acknowledgments

The authors would like to thank all the participants who took part in the study and Rachel van Besouw for her assistance in implementing the experimental procedure.

References

[1] J. Sundberg, *The science of singing,* Northern Illinois Press, Illinois (1987)

[2] P. Howes, J. Callagham, P. David, D. Kenny, W. Thorpe, “The Relationship Between Measured Vibrato Characteristics and Perception in Western Operatic Singing”, *J. Voice* 18(2), 216-230 (2004)

[3] J. Diaz, H.Rothman, “Acoustical Comparison Between Smaples of Good and Poor Vibrato in Singers”, *J. Voice* 17(2), 179-184(2003)

[4] H. Daffern, *Distinguishing characteristics of vocal techniques in the specialist performance of early music,* PhD Dissertation, University of York (2008)

[5] J. Bretos, J. Sundberg, “Measurements of Vibrato Parameters in Long Sustained Crescendo Notes as Sung by Ten Sopranos”, *J. Voice* 17(3),343-352 (2003)

[6] I. Ferrante, “Vibrato rate and extent in soprano voice: A survey on one century of singing”, *J. Acous. Soc. Am.* 130(3), 1683-1688 (2011)

[7] E. Prame, “Measurements of the vibrato rate of ten singers”, *J. Acous. Soc. Am.* 96(4), 1979-1688 (1994)

[8] E. Prame, “Vibrato extent and intonation in professional Western lyric singing”, *J. Acous. Soc. Am.* 102(1), 616-621 (1997)

[9] R. van Besouw, D. Howard, “Effect of carrier on the pitch of long-duration vibrato tones”, *Proceedings of 9th International Conference on Music Perception and Cognition,* 851-859 (2006)

[10] R. van Besouw, D. Howard, “Range of tuning for tones with and without vibrato”, *Music perception* 26(2), 145-155 (2008)

[11] V. Verfaille, C. Guastavino, P. Depalle, “Perceptual evaluation of vibrato models”, *Proceedings of the Conference on Interdisciplinary Musicology* (2005)

[12] J. Sundberg, “Effects of the vibrato and the ‘singing formant’ on pitch”, *Musicologica Slovaca* 6, 51-69 (1978)

[13] J. Shonle, K. Horan, “The pitch of vibrato tones”, *J. Acous. Soc. Am.* 67(1), 246-252 (1980)

[14] C. d’Alessandro, M. Castellengo, “The pitch of short-duration vibrato tones”, *J. Acous. Soc. Am.* 95(3),1617-1630 (1994)

[15] H. Gockel, B. Moore, R. Carlyon, “Influence of rate of change of frequency on the overall pitch of frequency-modulated tones”, *J. Acous. Soc. Am.* 109(2), 701-712 (2001)

[16] J. Brown, K. Vaughn, “Pitch center of stringed instrument vibrato tones”, *J. Acous. Soc. Am.* 100(3), 1728-1735 (1996)