

Variability of Vibrato

-A Comparative Study between Japanese Traditional Singing and Bel Canto-

Kenji Kojima*, Masuzo Yanagida* & Ichiro Nakayama**

*Department of Knowledge Engineering, Doshisha University, Japan

**Department of Musicology, Osaka University of Arts

dtc0708@mail4.doshisha.ac.jp, myanagid@mail.doshisha.ac.jp, nkym1@gold.ocn.ne.jp

Abstract

Several styles of vibrato are investigated, comparing those of Japanese traditional singing with western bel canto from viewpoints of stationarity, depth, rate, build-up time, and synchronization between F_0 and power. Analyzed as Japanese traditional singing are Noh, Kyogen, Heikyoku(Biwa), Shomyo, Kabuki and Nagauta. Singers of Japanese traditional singing include living national treasures. Data are taken from a CD database edited by Nakayama, the last author of this report.

1. Introduction

Nakayama [1] made recording of Japanese traditional singing and the recorded data were published in the end of 2002 as a database consisting of 18 CDs with a detailed guide. Before completion of recording, the authors began investigating Japanese traditional singing. They investigated Japanese traditional singing from a viewpoint of vowel characteristics and vibrato comparing with western bel canto [2] putting emphasis on vowel characteristics. It was revealed that formant frequencies characteristic to genres are observed both in bel canto and Japanese traditional singing and that tendencies of formant shift within a singer from natural utterance to singing utterance are somewhat different depending on music genres of the singers. Vibrato, periodic change of fundamental frequency F_0 , was investigated just from a viewpoint of its depth and required time for build-up.

Tastes in music may differ depending on culture and also on individuals. By singing a song, the "pitch" (fundamental frequency) and the "note value" (duration on score) of each note is usually prescribed in scores in case of western music except a few modern songs like "Pierrot Lunair" by Schönberg, in which singers are specified to change the pitch to any other height as soon as she reaches the specified pitch. Temporal intensity of each note is also roughly specified usually by 9 levels, from *ppp*(pianississimo) to *fff*(fortississimo), though the intensity levels are not quantitatively defined, but just subjective to singers. So, the prosody, defined as temporal change in F_0 and intensity, in song is regarded as almost specified beforehand and only a few freedom is left for singers. In that sense, prosody in songs is not spontaneous but highly controlled under strict specifications, however, there is freedom left for singers. Vibrato is one of the freedom left for singers.

Singers have freedom in vibrato, that results in varieties of personal vibrato styles. However, there might be certain tendencies regarded as characteristic to genres. The previous report [2] describes gross features of vibrato in Japanese traditional singing together with those of bel canto, without showing any examples. Details are described in this report showing con-

crete examples of vibrato.

In western classical singing, "vibrato" or "non-vibrato" is often specified on score, as appropriate vibrato is usually preferable for long notes, but its depth and rate are usually left for singers' freedom. Also in Japanese traditional singing, vibrato and melisma (melody on a single syllable in letters) are specified by putting special marks on letter scores. Preference for features of vibrato, however, differs depending on regional culture and taste of individuals. For example excess vibrato and too fast (faster than about 10Hz) vibrato are usually not accepted in western classical singing. We can recognize the differences of cultural preferences even by simply listening to the songs. Investigated here are stationarity, depth, rate, required time for build-up and synchronization between F_0 and power of vibrato in Japanese traditional singing compared with those in bel canto.

2. Variability of vibrato

Some examples of vibrato in Japanese traditional singing are shown in Fig. 1, where the abscissa represents the time and the ordinate, F_0 . Depicted in Fig. 1 are "Noh", "Kyogen", "Biwa(Heikyoku)", "Shomyo", "Kabuki" and "Nagauta", from the above-mentioned CD database [1]. F_0 pattern of Noh begins with a spike-like sharp change, passing through a constant pitch part, and reaches a stable vibrato of almost perfect fourth wide. Shomyo also shows a similar pattern though the starting part is not so sharp. Kabuki also shows a similar pattern though it lacks the sharp change at the beginning on F_0 patterns, and the vibrato depth in the last half is only a semitone. The rest three show stable vibrato of a semitone or a whole tone wide starting just after the beginning of voicing.

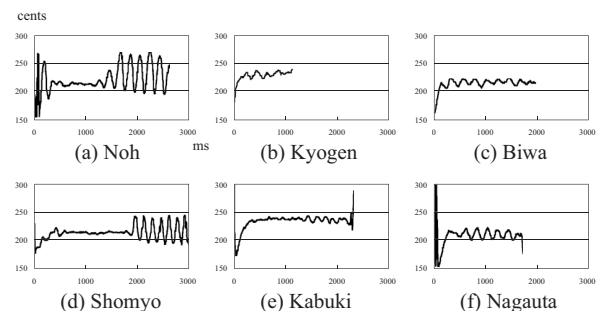


Figure 1: Examples of vibrato in Japanese traditional singing. Abscissa: time[ms], ordinate: F_0 [Hz].

Though above-mentioned features of vibrato are described

as if those were characteristic to genres, those might not be the reality because the number of informants for each genre is not so large, but those features might be characteristic to individual singers.

A typical vibrato pattern of bel canto is separately shown in Fig. 2, where vibrato builds up starting at the beginning of voicing reaching a stable vibrato of a whole tone wide. Vibrato depth and vibrato rate commented in the figure will be explained in section 3.

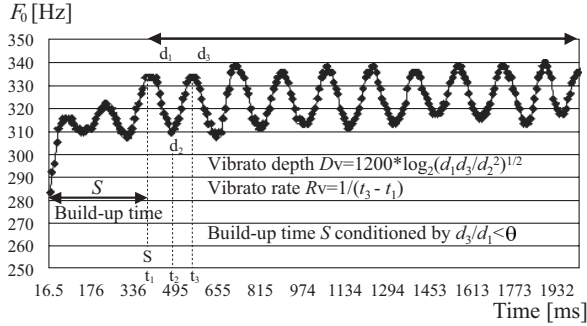


Figure 2: A typical vibrato pattern of bel canto. Vibrato depth is defined as $1200 * \log_2(\sqrt{d_1 * d_3}/d_2)$ [cents] and vibrato rate as $1/(t_3 - t_1)$ [Hz].

3. Vibrato depth

As observed in Fig. 1 some of Japanese song genres, such as Noh and Shomyo, show a much wider range of change in F_0 than bel canto. The ratio of the highest F_0 to the lowest F_0 within a period reaches 4/3, or perfect fourth higher, for Japanese traditional singing, while that of bel canto is only about 9/8, or one whole tone, though some other Japanese genres show very shallow vibrato. To investigate the situation more precisely, instantaneous “vibrato depth” at $t = t_2$ in Fig. 2 is defined as $1200 * \log_2(\sqrt{d_1 * d_3}/d_2)$ [cents], expressing the width of change in F_0 within a period. “Cent” is a unit of measuring frequency ratio in a logarithmic scale defined as an octave corresponds to 1200 cents or one semitone corresponds to 100 cents.

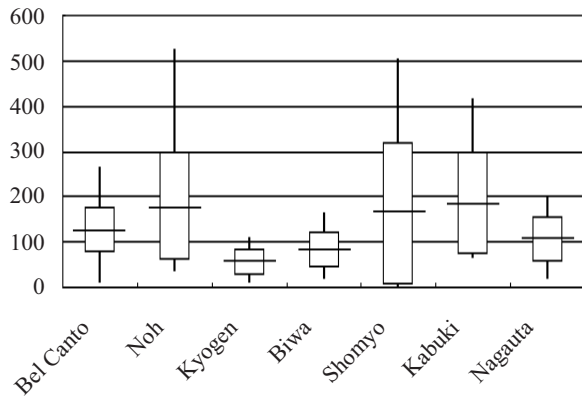


Figure 3: Average depth[Hz], standard deviation[Hz], maximum and minimum depths[Hz] of vibrato for each genre including bel canto for comparison.

Figure 3 shows average vibrato depth, its standard deviation, maximum and minimum values for each genre including bel canto for comparison [2]. From Fig. 3 we can recognize that the maximum depth of Japanese traditional singing is 500 cents, i.e. perfect fourth, while that of bel canto is much smaller than 300 cents, i.e. three semitones. In Japanese singing, perfect fourth may have important role as it is used as the interval between the two voice parts in “Shomyo”, a chorus by Buddhist bonzes at formal ceremonies. Perfect fourth is observed also as melodic jump at chanting “sutras(Buddist scriptures)” and as musical interval between two parts in “sutra”.

4. Vibrato rate

Instantaneous “vibrato rate” at $t = t_2$ is defined as $1/(t_3 - t_1)$ [Hz], expressing the average rate of change in F_0 from t_1 to t_3 . Figure 4 shows average vibrato rate, its standard deviation, the maximum and minimum values for each genre including bel canto for comparison. From Fig. 4 we can see that the average vibrato rate is within the range from 4 to 5.5 Hz independent of its genre.

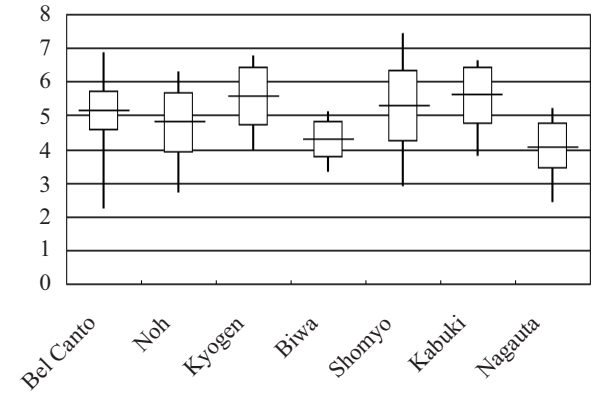


Figure 4: Average rate[Hz/s], standard deviation[Hz/s], maximum and minimum rates[Hz/s] of vibrato for each genre including bel canto for comparison.

5. Stationarity of Vibrato

By visual inspection or listening the sounds, every vibrato in bel canto seems to be stationary or stable from just after the beginning of voicing up to almost the end of utterance, while that of Japanese traditional singing does not always seem to be so stationary, or rather than that, as shown later in Fig. 6 (b), F_0 of Japanese traditional singing sometimes lacks vibrato even for a long vowel, or its curve often shows nonstationary changes as seen in Fig. 1 (a), (d) or (e).

Stationarity can be, in a sense, evaluated by variances in vibrato depth and vibrato rate. If the both the variances are small for a genre, it can be said that the vibrato of the genre is stationary. As mentioned above, the rate of periodic change in F_0 seems not so different from 5Hz for vibrato parts, that means variance in vibrato rate is small for both Japanese singing and bel canto. So, it would be said that stationarity can be roughly evaluated only by variance in vibrato depth. Allowing that interpretation, Fig. 3 says that Noh and Shomyo are nonstationary including deep vibrato and shallow vibrato as seen in Fig. 1 (a) and (d), while Biwa and Nagauta seem to be stationary keeping

rather stable vibrato from the beginning of voicing.

6. Build-up time of vibrato

We observe difference between Japanese singing and bel canto also in build-up time depicted as the point S in Fig. 2. S is defined as the time point at which F_0 forms the first peak after which d_3/d_1 , the ratio of adjacent local peaks in F_0 curve, is less than a threshold, 1.2 as a preliminary value here.

The situation is seen in Fig. 5, which shows that build-up time in Japanese traditional singing is usually longer than bel canto. Noh, Shomyo and Kabuki, in particular, show average values longer than twice the average value of bel canto, with extremely long build-up time as the longest value.

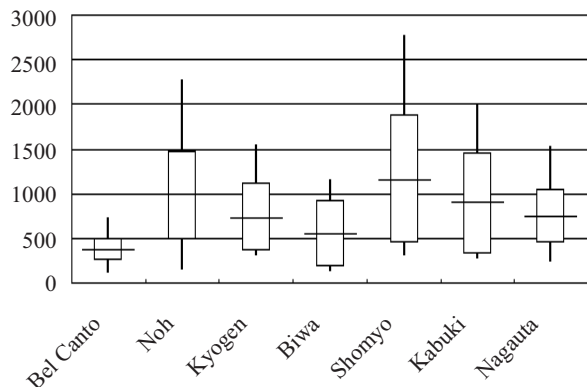


Figure 5: Comparison in build-up time[ms] between Japanese singing and bel canto.

7. Synchronization between F_0 and short-term power

7.1. Conceptual consideration

Synchronization between F_0 and power is an important factor of vibrato, because perceptual feeling of vibrato is formed by periodic changes of both F_0 and power. In case the signal in concern is a single component sound, its vibrato is simply considered as perceptible sinusoidal change of the frequency with a rate below about 20Hz. In that case power can be kept constant. Apart from vibrato, sinusoidal change of power causes sensation of amplitude modulation or “beat” provided that the modulation rate is below about 20Hz. In case of sounds having harmonic structure like vowels, however, complexities arise from the fact that temporal change of F_0 produces change of power in a complicated manner because of the reasons described below:

1. Vowel spectra consist of equally spaced harmonic components with a frequency interval corresponding to F_0 as shown with vertical solid lines in Fig. 7.
2. Spectral envelopes of vowels are multi-peaked shapes having ascendant slopes and descendant slopes along the frequency axis as shown with the envelope in Fig. 7.
3. In case F_0 varies, amplitudes of harmonic components varies along the spectral envelope as shown with arrows in Fig. 7, showing the situation where F_0 increases by Δ .
4. The total power is determined by summing up the power of all components.

5. It is natural to think that amplitude of vocal chord vibration is not kept constant, but is necessarily synchronized with the change of vocal chord vibration.

Taking issues 2 through 4 noted above into consideration, it is obvious that temporal variation of F_0 causes certain change in power, but we don’t know if the power increases or decreases, as it depends on the spectral shape and F_0 . To make the matter more complicated, issue 5 says that there must be a temporal change in glottal excitation power, which is supposed to be necessarily and physically synchronized to the change of F_0 anyway in order to give a periodic change to F_0 , or glottal excitation. Let’s call the former “power change by envelope” and the latter, “power change by source”.

Power change by source is supposed to be larger than power change by envelope and good singers are supposed to be able to control the power change by source. The power change by envelope seems to be difficult to control, because harmonic components may be located on ascendant slope or descendant slope and it is difficult to control every component by a single variable F_0 . Good singers, however, may be able to control it by adjusting formant frequencies, or spectral peak positions. Precise and detailed measurements are required to confirm this issue.

Anyway, power change occurs if F_0 changes, so in case F_0 shows a periodic change, so, the power, synchronously either in phase or in reverse phase. It is inevitable because it is impossible to cancel or compensate the “power change by envelope” by controlling the source power or the amplitude of glottal excitation. On the contrary, fluctuation or periodic change only in power is possible by amplitude modification or periodic change of excitation amplitude keeping F_0 constant, though it may be somewhat difficult to change excitation amplitude keeping the excitation frequency constant. An actual example of this type of singing is shown in Fig. 6 (b).

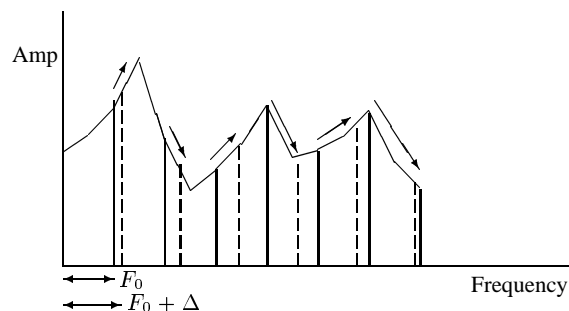


Figure 7: Change in F_0 on a harmonic structure.

7.2. Observed phenomena

Fig. 6 compares (a) bel canto, (b) Shomyo and (c) Noh by waveforms of vowels(top), F_0 (second top), short-term power and mid-term power(second bottom) and normalized short-term power(bottom), where “short-term power” means squared-sum of Hanning windowed waveform of 20ms long and “mid-term power” means that of 100ms and “Normalized short-term power” means short-term power normalized by mid-term power.

It is recognized that vibrato in bel canto, tenor in this case, builds up rapidly, as mentioned in section 6, and shows stationary temporal patterns both on F_0 and power. Figure 6 (b) is an interesting example observed in “Shomyo”. The informant of this data is Shuin Matsukubo, the former head of famous temple

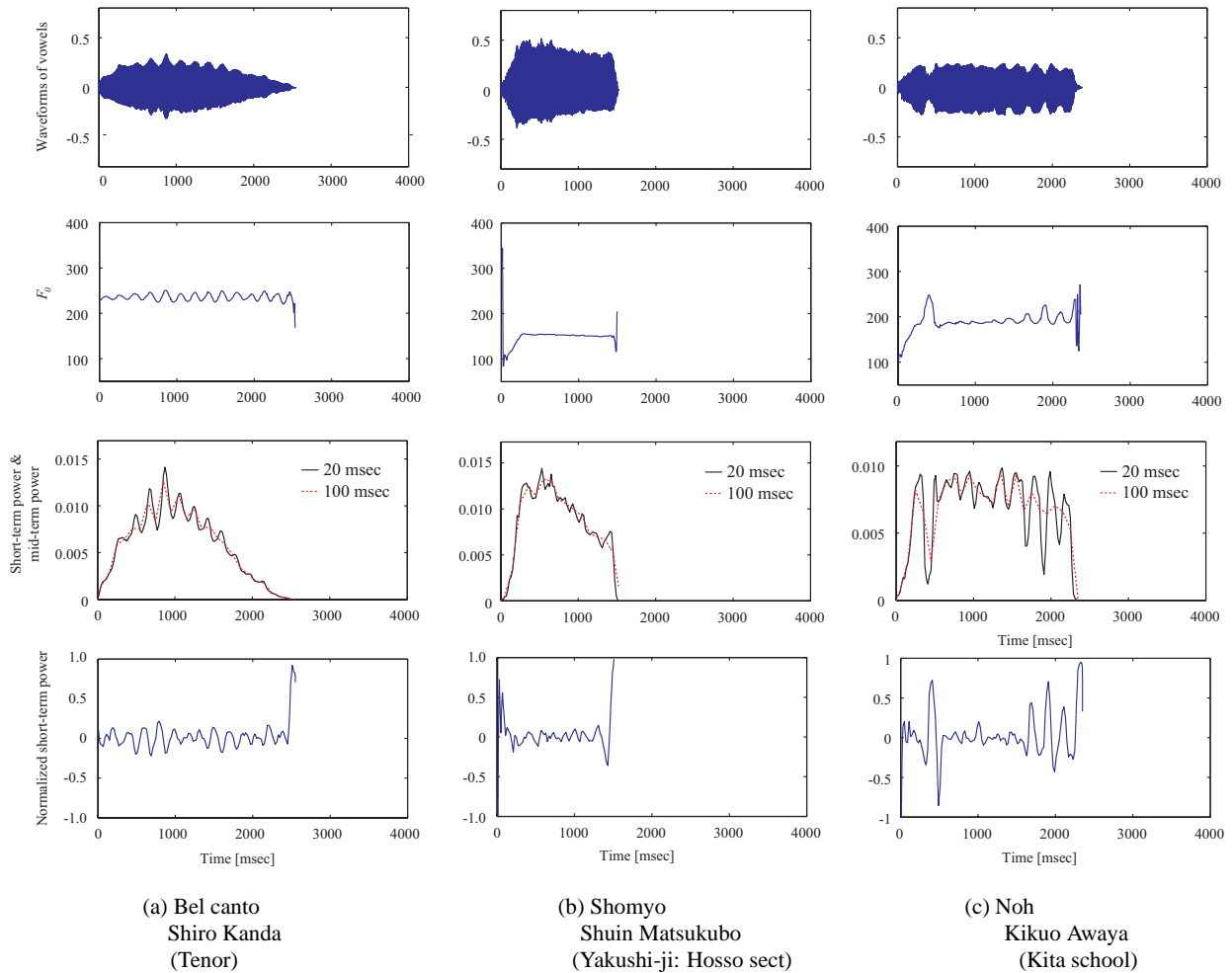


Figure 6: Comparison of synchronization between F_0 and power.

“Yakushi-ji”, belonging to Hosso sect. No vibrato is observed in F_0 , but temporal variation is observed on power though periodicity is not clear. It is a kind of amplitude modulation, discussed in the previous section, but not vibrato.

Figure 6 (c) is a vowel of “Noh”. Its informant is Kikuo Awatani, living national treasure for Kita school. In this case one strong peak is recognized on F_0 pattern just after build-up and then F_0 is kept rather flat for a while, then strong but non-stationary vibrato emerges in the ending part of the utterance. Temporal power variation is very wide for this case reaching 80%, though vibrato depth in F_0 is about perfect fourth.

8. Conclusion

Vibrato of several genres of Japanese traditional singing is compared with that of western bel canto. It appears to be much deeper and takes much longer time to reach a steady state than western bel canto, though there seems no significant difference in vibrato rate. Vibrato of bel canto builds up rapidly and shows stationary change in both F_0 and power, while vibrato of Japanese traditional singing shows diverse appearances such as stationary one like that of bel canto, time variant type with a spike-like change in the beginning with mid part of little change

having deep vibrato in the ending part. There can be AM-type singing as an extraordinary example in Shomyo. The maximum vibrato depth in Japanese traditional singing is perfect fourth, while that in bel canto is a whole tone. Synchronization between F_0 and power is discussed from a viewpoint of mutual relation between spectral envelopes and the harmonic structure of vowels, though detailed investigation on actual data is left for future research.

This research was supported by Academic Frontier Project promoted by Doshisha University and by Grant-in-Aid for Scientific Research on Priority Areas, MEXT.

9. References

- [1] Nakayama, I.(ed.), 2002. Singing, Chanting and Reciting in Japanese(DB of 18 CD set), Osaka: Ad Popolo.
- [2] Yanagida, M.; Nakayama, I., 2002. Comparative studies on Japanese traditional singing and bel canto. *Forum Acusticum Sevilla*. España, MUS-03-002-IP.