Voice Production Mechanisms of Vibrato in Noh

Ikuyo Yoshinaga¹, Jiangping Kong¹,²

¹ Department of Chinese Language and Literature, Peking University, China
² Center for Chinese Linguistics, Peking University, China

beibe.ily@gmail.com, jpkong@pku.edu.cn

Abstract
Vibrato used in Noh play was investigated using electroglottographic and acoustical analyses. Laryngeal movements were successfully obtained from the EGG signal and its derivative in order to study how the peculiar and expressive voice qualities were produced during vibrato. The slow wide vibrato of Noh was achieved not only by the F0 modulation but also by changing the degree of constriction at the vocal folds and the supraglottal structures. These complicated movements at the end of the resonator seemed to add color and variety to the voice quality, and produce an extremely unique vibrato.

Index Terms: vibrato, Noh, supraglottal constriction

1. Introduction

Noh is a Japanese traditional performing art which has an excellent reputation for its solemn performance style and its unique voice quality in singing. Shite (leading role), who is wearing a mask, sings and dances in slow mode without showing any facial expression. From the oral arts viewpoint, the singing techniques in Noh was influenced by Shomyo (sabda-vidya, in Sanskrit), as it has been called, was one of the five fields of academic study in ancient India.

Research on the singing voice of Noh is quite rare. Regarding the Noh voice quality, [1] suggests, ‘strong interaction between vocal tract resonance and vocal fold vibration seems to play a major role in producing these expressive voice qualities.’

Vibrato is a technique not only used in the human voice but also in various musical instruments, research on the vibrato in Western operatic singing revealed that a vibrato with the rate of between 5.5 and 7.5 undulations per second and the extent of ±1 or 2 semitones is considered to be a good vibrato [2, 3].

Research on phonation types using electroglottography (EGG) has identified the vibratory phase characteristics and corresponding laryngeal configurations of a series of phonation types common in current phonetic usage [4-6]. Recent investigation using high-speed camera, X-ray, Kymography, and other methods has revealed ventricular and aryepiglottic fold oscillations in Asian vocal cultures and in some ethnic and pop music [7-9]. And pressed, vocal-ventricular mode (VVM), and growl voice have been estimated to be used in Noh singing using electroglottographic and acoustical analyses [10].

This paper describes the laryngeal movements during vibrato in Noh play and the reflective acoustical characteristics.

2. Methods and materials

This section describes the calculation method of EGG and DEGG-based parameters and voice materials.

2.1. Parameter calculation method

The time-based parameters of the EGG and DEGG waveforms also yield information about time patterns of the vibratory events. Three EGG-based parameters, F0, OQegg, and SQegg, are defined as: F0=1/period, OQegg%=de-contact phase/period*100 and SQegg%=de-contacting/contacting*100. The derivative of the EGG signal (DEGG) is employed in this study, in which the glottal closing instance (GCI) and glottal opening instance (GOI) are determined as the maximum and minimum values of the DEGG waveform (See Fig. 1).

![EGG waveform, DEGG waveform and phases of vocal fold contact.](image-url)
Triple GCIs often appear when F0 is under F3 (175 Hz) (See Fig.2). Three DEGG-based parameters are extracted at most using our novel analysis methods: $t_0$, $t_1$ and $t_2$. Their definitions are described as follows: $t_0=GCl(n+1)-GCl(n)$, $t_1=GCl_2(n)-GCl(n)$ and $t_2=GCl_3(n)-GCl(n)$, ‘n’ is the number of the glottal cycles. A phase difference between GCI and GCI$_2$ is described as $(t_1/t_0)*360°$, and that between GCI and GCI$_3$ is described as $(t_2/t_0)*360°$.

![Figure 2: Definition of $t_0$, $t_1$, and $t_2$](image)

2.2. Voice material

The Noh play “Tsurukame” sung by a successor of Yoshio Hosho, a living national treasure, was recorded at the Nohgaku stage in Tokyo. The EGG signal was obtained by an EGG system (Electroglottograph Model 6103; Kay, USA). The audio signal was recorded by a Sony Electret Condenser Microphone. Those signals were simultaneously recorded and digitized at 16-bit resolution at a sampling frequency of 44.1 kHz. The data processing was performed by Matlab-based program, which was developed by the authors.

3. Electroglottographic analysis

This section describes phonatory characteristics of the vibrato using EGG parameter and phase analyses.

3.1. EGG parameter

Fig. 3 shows audio and EGG signals of five vibrato cycles. Period doubling is seen in the center of the selected EGG, it suggests that the vibrato is achieved by the combination of F0 modulation and the switching of phonation types. Pressed voice and VVM appear alternately in this example.

![Figure 3: Audio and EGG signals of a vibrato /ba/ from the phrase ‘haru ni nareba’ and the selected EGG signal](image)

Fig. 4 shows F0, OQ$_{egg}$ and SQ$_{egg}$ of a vowel /o/ from the phrase ‘kin gin no’. OQ$_{egg}$ curve appears as the same phase as F0 contour, however, SQ$_{egg}$ curve appears as the opposing phase as F0 contour. The mean F0 of 200 glottal cycles is 186 Hz, that of OQ$_{egg}$ is 51%, and that of SQ$_{egg}$ is 273%. The differences between the minimum and maximum values of OQ$_{egg}$ and SQ$_{egg}$ up to 35% and 508% respectively, while that of F0 remains at approximately 70 Hz (±3 semitones). The extreme change of parameter values undoubtedly results from the switching of phonation types.

![Figure 4: F0, OQ$_{egg}$ and SQ$_{egg}$ curves of a vibrato /o/ from the phrase ‘kin gin no’](image)

3.2. Phase difference between GCIs

Fig. 5 shows F0 contour and a phase difference between GCI and GCI$_2$ in a vibrato /o/ from the phrase ‘sikitae no o’. The mean F0 is approximately 186 Hz, GCI$_2$ appears only in the low F0 region of the vibrato cycle with approximately 190° phase after the GCI. It is hypothesized that the appearance of GCI$_2$ is reflective of the ventricular fold adduction. The earlier researches on VVM using the high-speed videendoscopy revealed that the ventricular closure occurred during the 480° - 560° interval during the vocal folds were open [7]. In other words, the ventricular closure occurred during the 120°-200° interval every other glottal cycle because the frequency of the ventricular fold oscillation was F0/2 in the above case. From the EGG waveshape viewpoint, period doubling in the EGG waveform represents VVM [5]. The ventricular fold constriction is estimated to occur at average of 171° phase difference after the glottal closure in Tibetan chants [11]. It seems plausible that GCI$_2$, which appears around 190° phase after the GCI in this case, is judged as the closing instance of the ventricular fold oscillation.
4. **Acoustical analysis**

In this section, acoustical parameters are studied to reveal how irregular movements at the laryngeal position are reflected in spectral domain. Mainly formant frequencies and the contours of F0 and intensity are discussed.

4.1. **Formant frequencies**

Fig. 7 illustrates F0, intensity, F2 and F1 curves of a vibrato /o/ from the phrase ‘sikitae no o’. Basically, F1 and F2 curves undulate in phase, and F0 and formant frequencies vary in opposite phase. The fluctuations of F1 and F2 reach approximately 200 Hz and 150 Hz. There are three factors that seem to cause formant frequencies to fluctuate: 1) by manipulating the throat muscles to move the larynx rapidly up and down to create fluctuations in pitch; 2) by the supraglottal constriction which changes the length of the vocal tract; 3) by periodically changing the constriction level of the vocal folds and that of the supraglottal structures. The factor 1 and 2 contribute to shorten the vocal tract in the low F0 region. However, the factor 3 is likely to be the main reason for the rise in formant frequency because according to perturbation theory, the constriction of the vocal tract near a point of maximum pressure (node) raises the formant frequency [12].

4.2. **F0 and intensity contours**

Vibrato of Noh is characterized with the rate of ±5 undulations per second and ±3 or 4 semitones. The vibrato, characterized with a slow rate and a large extent, reflects a slow and solemn performance style of Noh.
Fig. 8 shows the narrow-band spectrogram and the intensity contour of a vibrato /ba/ from the phrase ‘haru ni nareba’. The intensity in a vibrato depends mainly on the strongest spectrum partial, which normally is the partial lying closest to F1, three types of the intensity contour are proposed: 1) the same phase as F0; 2) the opposite phase as F0; 3) twice as fast as the vibrato frequency [2, 3]. However, besides these three types, the intensity contour which is three times as fast as the vibrato frequency is observed in the Noh (see Fig.8). Three intensity peaks, which are indicated by gray arrows in Fig.8, occur during one vibrato cycle. The reason for its occurrence is that the closest partial to F1 changes according to the F1 fluctuation in the case of the vibrato in Noh, because F0 and F1 appear in opposite phase. These kinds of complicated intensity contours also seem to contribute to produce unique voice qualities of Noh.

5. Conclusions

Besides F0 modulation, the supraglottal constrictions were involved in the voice production during vibrato. Period doubling in the EGG waveform, dramatic fluctuation of EGG parameter values, and double or triple GCIs in the DEGG waveform were supportive evidences for the supraglottal constrictions. Though the vibrato of Noh is very slow and wide, it does not sound dull, it sounds rather impressive because the supraglottal constrictions add an effective modulation to it, and produce a special voice quality.

There are two differences between the supraglottal constrictions in the Noh voice and that in the creaky voice: 1) the latter only occurs with very low pitch, however, the former occurs not only in the low-pitch region but also in the middle-pitch region; 2) the vocal folds oscillate regularly in the former, however, they oscillate irregularly in the latter.

6. Acknowledgements

The authors thank Yoshio Hosho of the living national treasure from the Hosho school of Noh, his pupils and Mr. and Mrs. Ueda from Toshiba for the recordings and valuable advice on this study.

This study was supported by China Social Sciences Funds (Grant No: 10&ZD125).

7. References