ACOUSTIC CHARACTERISTICS OF KOREAN TRADITIONAL SINGING VOICE: A PRELIMINARY REPORT

Seung-Jae Moon
Ajou University
Suwon, Korea

ABSTRACT

Most Koreans agree that Korean traditional singing voice has a very peculiar sound comparing to Western singing voice.

The goal of this paper is to investigate the acoustic characteristics of Korean traditional singing voice called 'Pansori'. Materials are analyzed from 3 male professional singers and 4 female professional singers. Their singing was compared with their own conversation and other non-singers' conversation.

Long term average spectra indicated that all the singers showed a much less spectral tilt than non-singers. The phenomenon was prevailing for professional singers not only in their singing, but also in their conversation. This suggests that it is not the result of a temporary effort but it may involve a certain permanent change in their physiological configuration. (To assess this hypothesis, voice source should be looked at directly. Therefore, in further research, using Rothenberg mask (Rothenberg, 1973) is strongly recommended.)

In addition to LTA, individual vowel formants will be studied later.

1. Problem

It is unanimously agreed, at least among Koreans, that Korean traditional singing voice called 'Pansori' sounds quite different from Western singing voice. One can see jugular veins sticking out clearly when a professional singer vocalizes. Though the vocalization itself seems very strenuous, a well trained singer can, and is required to perform more than 8 hours in stretch for one whole performance.

It takes years of training to become a respected singer who 'achieved the sound'. When one is considered to have achieved the sound, he is called a 'great singer,' 'Myung-Chang'.

While Western singing voice has been extensively studied (Sundberg, 1987), Korean traditional singing has not been an object of a scientific research. It has been mainly in the realm of subjective judgement and appreciation.

The goal of this paper is to investigate the acoustic characteristics of the Korean traditional singing voice.

How is it different from Western vocalization? Sundberg showed the so-called singing formants in a professional singer's vocalization in singing. Can we find that kind of special modification in Korean traditional singing too?

It is said that many famous singers practiced in a special, natural environment such as in front of a waterfall, in windy woods, in rain, or on beach. Why? If we know something about those favored environmental sounds, could we infer the characteristics of the traditional singing voice from that?

These are the questions which will be raised in this paper.

2. Procedures

2.1 Subjects

The subjects for this study were three male professional singers (S, T, P) and three female professional singers (K, A, C) all of whom are considered to be 'Myung-Chang's. Also one female professional singer (O) was studied who did not yet make a 'Myung-Chang'.

2.2 Materials

For each subject, recordings of both singing and regular conversation were obtained whenever possible. It was for the purpose of comparing his/her singing with conversation. However, for subject K, obtaining conversational material was not possible since she had passed away some time ago.

2.3 Data Acquisition

Some of the recordings were made using Panasonic Hi-Fi VHS VCR (model PV-4351) and others were made with TASCOM DA-P1 DAT. Materials recorded with the VCR were obtained from TV broadcasting. Other materials recorded with the DAT were from Munhwa Broadcasting Company's reference library.

These recordings were digitized using Kay CSL 4300B at 20kHz sampling rate. Input level was manually adjusted in such a way that the vocal part of the recordings was not overloaded. (In Pansori, there was a drum accompaniment with the singing which was sometimes quite loud comparing to the singing itself. The drum sound was overloaded when, otherwise, the vocal part was too small. Then the overloaded drum sound was manually deleted later.)

2.4 Analysis

For each recording, long term average spectra (LTA) were computed with the frame size of 25.6ms over several minutes of recording.

Then for each individual vowel formant measurements, spectrogram, FFT with 25.6ms window and LPC with the 20ms window were used. Pre-emphasis for LPC and FFT, which is usually employed in speech analysis, was set to 0 for the present study. For male singers, spectrogram filter bandwidth was set to 293 Hz or 391 Hz and for females, to 586 Hz.

3. Results & Discussion

We will start with the LTA's of some of the environmental sounds in which professional singers choose to practice their voice. Figures 1 and 2 show the LTA of a waterfall and of surf on the beach respectively.

In these figures, x-axis indicates frequency upto 8 kHz. Y-axis shows

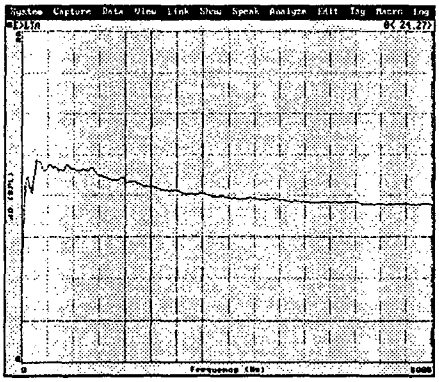

Figure 1

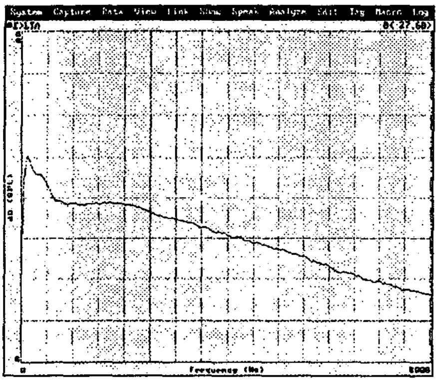

Figure 2

an arbitrary intensity scale.

Both of them share a common characteristic: that is, both of them show a very small tilt of energy as frequency goes up. This is especially true for the waterfall; it is practically similar to white noise. Even though we lack data of the sound from windy woods or raining forests, it seems plausible to assume that they too will show the similar pattern: that is, smaller tilt than usual human speech sounds, which typically exhibit -6dB/oct slope in a spectrum (Fant, 1960; Sundberg, 1987) as shown in Figure 3.

From this observation, it seems reasonable to hypothesize that one of the goals the singers try to achieve in their voice may be the lower degree of spectral tilt.

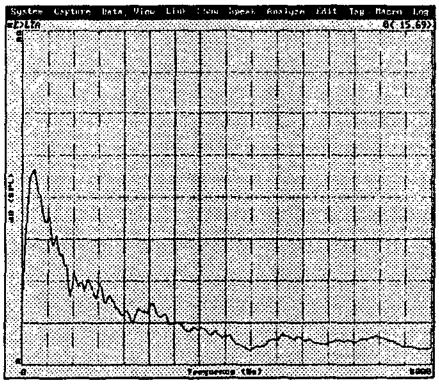

Figure 3

LTA's from six subjects show that their spectral tilt is indeed much smaller than normal voice. Only two are shown in the following figures (Figures 4 and 5 show the results from subjects S, and P respectively.)

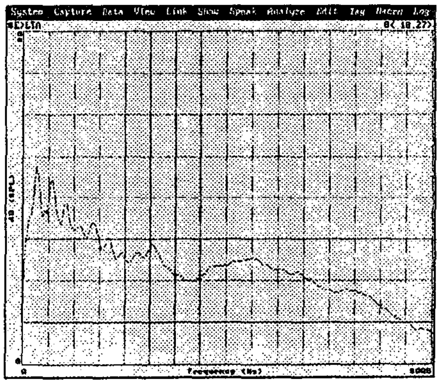

Figure 4

Upon examining the figures, particular attention should be paid to the relative value of amplitude, not the absolute value since input level could have been varied during the digitization.

Also notice the big prominent bumps between 2.5 kHz and 4 kHz in subject P. Another subject K also exhibits the same pattern. So far, it has been

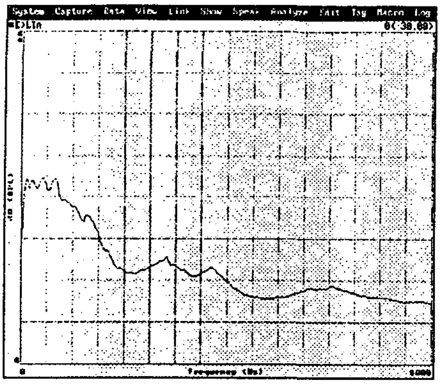

Figure 5

shown that professional Korean traditional singers adjust their spectral tilt to a degree much lower than -6dB/oct.

Is this the result of a temporary effort when they sing? Or is this the result of a permanent physiological change originated from extensive training? To shed some light on the question, regular conversational speech from each subject is compared with the singing voice. For illustration, the same subjects S (Figure 6) and P (Figure 7) are shown here. In these figures, LTA's from their conversational speech are superimposed on the LTA's from their singing (which are exactly Figures 4 and 5) to make comparison easier.

As we can see, spectrum envelopes of singing voice and regular conversation match almost exactly. In Figure 7, the difference comes from the different level of amplitude. If we slide the bottom line a little bit up, it will coincide nicely with the top line.

This is the case for all subjects except for one subject O. Figure 8 shows the results from subject O. The envelope on top is from singing, and

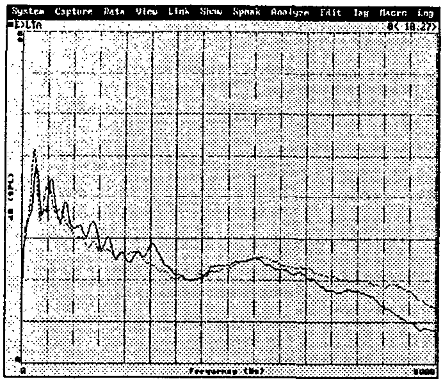

Figure 6

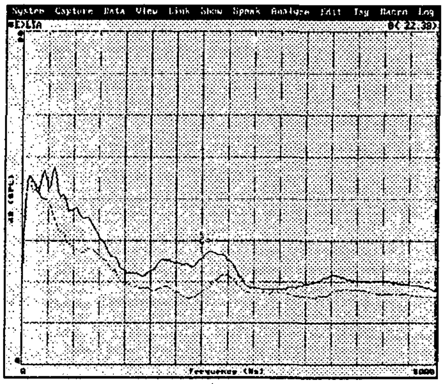

Figure 7

the bottom one is from conversation. Notice the lower frequency region. In this particular diagram, lower region's amplitude coincide with each other.

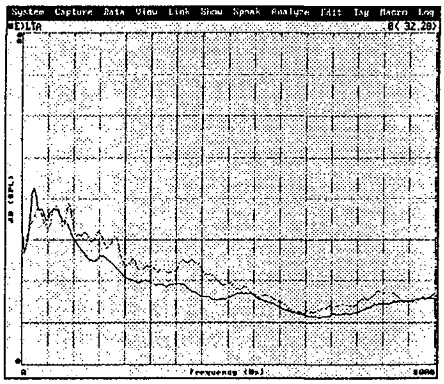

Figure 8

But in high frequency region, it is clear that the subject's singing voice energy is less attenuated than conversation. It is interesting to note that this is the subject who is in her youth and has not earned the title 'Myung-Chang' yet.

Based on this observation, we could assume that those spectral tilt changes observed in the professional singers are the results of permanent change in the voice source. Can we then say that in subject O's case, she is not yet in the state where the necessary physical change in the voice source has been completed.

4. Limitations

If this is the case, it seems clear that we need another kind of research to understand the characteristics of Korean traditional singing voice. As we have seen, normal conversation of professional singers also shows the same characteristics of the singing voice. That implies that the voice source has been changed permanently.

Speech signals used in this study were obtained in a conventional way: i.e. through a microphone in front of the mouth. That signal is the result of combination of a voice source (vocal folds) and a filter (vocal tract). There is no way to separate the source from the filter in a conventional way. And the mystery of Korean traditional singing voice seems to lie in the source. Therefore we need take out the source information from the combined signal by utilizing a certain method such as inverse-filtering implemented in Rothenberg's mask (Rothenberg, 1973).

Of course, individual vowel formant structure should be studied also. LTA, after all, can just show us a general tendency. Whether those spectral tilt changes were due to 'singer's formant' as in Western singing can be answered by looking at individual vowels.

As a preliminary report, the present study is believed to shed some light on the direction of the future research of Korean traditional singing voice.

REFERENCES

Fant, G. (1960). *Acoustic Theory of Speech Production,* Mouton, The Hague.

Rothenberg, M. (1973). "A new inverse-filtering technique for deriving the glottal airflow waveform during voicing," *J. Acoust. Soc. Am.* 53: 1632-45.

Sundberg, J. (1987). *The Science of the Singing Voice.* Northern Illinois University Press, Dekalb, IL.

ACKNOWLEDGEMENT

This paper was supported by NON DIRECTED RESEARCH FUND, Korea Research Foundation.