



Longitudinal Changes of Selected Voice Source Parameters

Hideki Kasuya¹, Hajime Yoshida², Satoshi Ebihara³, Hiroki Mori¹

¹ Graduate School of Engineering, Utsunomiya University, Utsunomiya, Japan

² Yoshida Clinic, Nayoro, Hokkaido, Japan

³ Kyoundo Hospital, Tokyo, Japan

kasuyah@snow.ucatv.ne.jp, nort3381@north.hokkai.net, ebihara@po.kyoundo.jp, hiroki@speech-lab.org

Abstract

Longitudinal changes were investigated for selected voice source parameters: fundamental frequency (F_0), jitter (period perturbation quotient, PPQ), shimmer (amplitude perturbation quotient, APQ) and glottal noise (normalized noise energy, NNE). Acoustic analyses were made on the sustained phonation of the Japanese vowel /a/ of 20 males and 38 females with no laryngeal disease, which were recorded over periods ranging from 10 to 18 years. The longitudinal change of the parameters was evaluated by the statistical t-test, revealing that: 1) strong individuality existed for significant longitudinal changes in the acoustic parameters, 2) F_0 falling is a more typical tendency of vocal aging of females than males, while F_0 rising, which has been pointed out in the previous reports for males, was not found, 3) shimmer is a more observable indication of vocal aging than jitter, and 4) glottal noise in a high-frequency region tends to increase with aging.

Index Terms: vocal aging, aging voice, fundamental frequency, jitter, shimmer, glottal noise

1. Introduction

Acoustic aspects of vocal aging have long been of considerable research concern [1], because 1) normative acoustic data on vocal aging of healthy people would contribute to the diagnosis and effective therapy of voice disorders and 2) accumulation of knowledge on vocal aging is fundamental for basic speech science. The acoustic measurements dealt with so far consist of voicing source parameters, such as fundamental frequency (F_0), jitter, shimmer, pitch range, maximum phonational fundamental frequency range, speech intensity, intensity range, and so on, and vocal tract parameters, such as formant frequency and long-term average spectral shape. These measurements have been made on the basis of either cross-sectional [2-9] or longitudinal data [10-15]. In order to explore intraindividual vocal changes with aging, a longitudinal study would certainly be preferred to a cross-sectional investigation. However, it is quite demanding to continually acquire longitudinal vocal data over a long period, so vocal aging has usually been discussed primarily based on cross-sectional voice data. There are many factors that are responsible for the acoustic variations of voice with aging: anatomical and physiological alterations of the speech organs with aging, smoking habit, vocal load of subjects, and so on. From cross-sectional studies, therefore, it might be difficult to draw general conclusions about the vocal changes with aging without knowing how subjects are controlled in the experiment.

There have been a few longitudinal studies on vocal aging, mostly either comparative studies between archival voice data and the newly recorded voice of the same speaker [10-12, 14] or investigations based on vocal materials recorded over a relatively short period such as five years [13, 15]. In this paper, we investigate longitudinal changes in selected voice source parameters, i.e. fundamental frequency, jitter, shimmer, and

glottal noise, based on the voice data recorded over periods ranging from 10 to 18 years.

2. Method

2.1. Speakers and materials

We have performed a voice screening experiment for early detection of laryngeal cancer in northern Hokkaido, Japan, where no laryngologist is available [16, 17]. In the experiment, which has been carried out once a year since 1982, examinees pronounced "My name is *** and age is *** years old" in Japanese, followed by the sustained phonation of the Japanese vowel /a/ five times. The speech recordings of the sustained phonation were later analyzed acoustically and judged automatically whether the examinee should undergo a closer examination or not. We have already experienced more than 25,000 people in total. Some of the examinees have repeatedly participated in the screening experiment over the years and allowed us to accumulate their voice recordings.

We selected 20 males and 38 females with no laryngeal disease from all the examinees whose recordings were made at least five times over nine years or more. The average period over which the recordings were made was 13.4 years in males and 13.2 years in females.

2.2. Acoustic measures

We examined five acoustic parameters relevant to glottal flow, all of which were extracted from the sustained vowel phonations. These parameters were mean fundamental frequency F_0 , jitter parameter PPQ (fundamental period perturbation quotient), shimmer parameter APQ (amplitude perturbation quotient) and two types of additive glottal noise energy included in the speech signal, $NNEa$ and $NNEb$. PPQ and APQ used in this paper were originally proposed by Koike [18] and modified by the authors [16] to make the parameters robust against a slow F_0 change and are given as follows:

$$PPQ = \frac{1}{N-2k} \sum_{i=k+1}^{N-k} \left| \frac{p_i - \frac{1}{2k+1} \sum_{n=i-k}^{i+k} p_n}{\frac{1}{2k+1} \sum_{n=i-k}^{i+k} p_n} \right| \times 100 \quad (\%)$$

$$APQ = \frac{1}{N-2k} \sum_{i=k+1}^{N-k} \left| \frac{a_i - \frac{1}{2k+1} \sum_{n=i-k}^{i+k} a_n}{\frac{1}{2k+1} \sum_{n=i-k}^{i+k} a_n} \right| \times 100 \quad (\%)$$

where p_n is a period sequence, a_n an amplitude sequence, N the length of a sequence and $2k+1$ a window length. A sophisticated algorithm to detect both period and amplitude sequence

from period to period is given elsewhere [19]. In the experiment, the window length $2k+1$ was 3-point.

$NNEa$ and $NNEb$ were proposed by one of the authors and his colleagues and proved to be useful for the acoustic evaluation of laryngeal diseases [19]. A sustained vowel signal $s(n)$ is assumed to consist of a harmonic $h(n)$ and a glottal noise $g(n)$ component:

$$s(n) = h(n) + g(n)$$

$g(n)$ is estimated by a comb filter as

$$\hat{g}(n) = s(n) - \sum_{k=-L}^L a_k s(n - N_k)$$

where a_k is a normalized Hamming window coefficient and N_k is a period [19]. The normalized noise energy NNE is an acoustic measure of the amount of additive glottal noise energy included in the signal normalized by the entire speech energy expressed as a percentage and is defined as follows:

$$NNE = 100 \sum_n \hat{g}^2(n) / \sum_{k=-L}^L b_k^2 \sum_n s^2(n) \quad [\%]$$

where

$$b_k = -a_k, \quad k \neq 0 \\ = 1 - a_0, \quad k = 0$$

The first term of the denominator compensates for the noise energy mixed with the harmonic component, assuming the noise to be white. The amount of noise is evaluated both for the whole frequency range and for the frequency range higher than 1.5 kHz. $NNEa$ is for the whole range and $NNEb$ for the higher frequency range.

3. Results

Figure 1 illustrates examples of the longitudinal plots of $F0$, PPQ , APQ and $NNEb$ of a female subject, from the top, whose voice recording started at the age of 50 and ended at 63 years old. During this period, recordings were made seven times. The straight line in each panel is an estimated regression line. Whereas $F0$ falls significantly with aging ($p < 0.05$), APQ and $NNEb$ significantly increase as age advances ($p < 0.05$). The change of PPQ , however, is not statistically significant, meaning relatively stable vibration frequency of the vocal fold irrespective of aging.

Figure 2 is an example of a male subject, who yielded no statistically significant change with aging for the same four parameters, even at the 5% level.

Table 1 summarizes the result of the statistical t-test of the gradient of the regression line for 20 males. Of the 20 males analyzed, recording was started for three while in their 40s, six in their 50s, eight in their 60s, and three in their 70s. The numerals in each cell in the table indicate the number of subjects who showed a statistically significant change with aging. Numerals in parentheses are the number of subjects who showed a negative gradient, i.e. a fall or decrease of the acoustic parameter in question. Blank elements in the matrix signify that no one revealed significant changes. Two males (10%) who showed a significant $F0$ change both indicated negative gradients, i.e. $F0$ falling. No one showed significant changes for PPQ and only one subject (5%) resulted in a significant increase in $NNEa$. On the other hand, a significant increase was observed in five males (25%) for APQ and in seven males (35%) for $NNEb$.

Table 2 displays the same statistics as in Table 1 tabulated for the 38 female subjects. The number of statistically signifi-

cant subjects who showed a negative gradient, i.e. falling or decreasing tendency of a parameter, is again indicated in parentheses. Ten out of the 38 females (26%) showed significant $F0$ falling, and no rising change was observed. As in the case of males, PPQ and $NNEa$ showed fewer significant subjects than APQ and $NNEb$. Eleven subjects showed increasing APQ (29%) with aging and 14 males resulted in increasing $NNEb$ (37%). In PPQ , however, two females showed increasing changes but four indicated a decreasing tendency, meaning that there were vocal changes in only six subjects in total (19%).

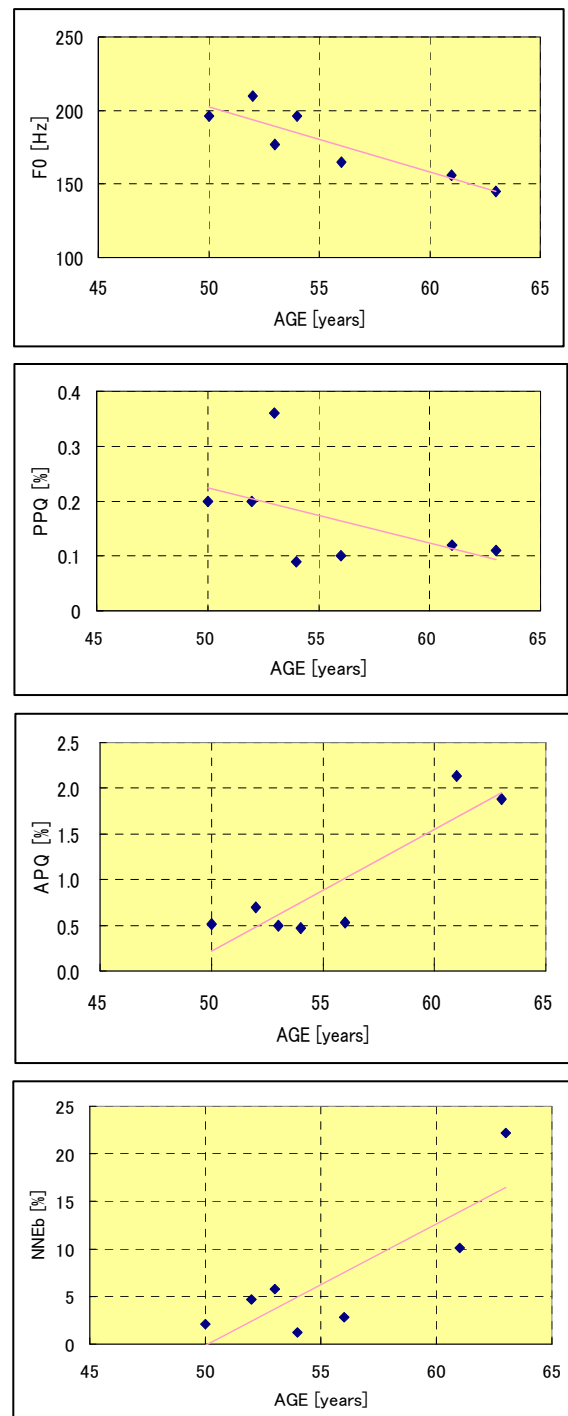


Fig. 1 Longitudinal plots of four acoustic parameters for a female subject, $F0$, PPQ , APQ and $NNEb$ from the top.

4. Discussion

It is obvious from Tables 1 and 2 that strong individuality of significant longitudinal changes exists in the acoustic parameters; the maximum rate of change in males is 35% for *NNEb*, and that in females is 37% for the same parameter, indicating that in just over 60% of the subjects, the vocal parameters do not change significantly with aging but rather time-to-time variations of the parameter values are dominant. However, the statistics of the vocal parameters obtained still deserve to be discussed in terms of the longitudinal change.

As for the fundamental frequency, a large number of investigations have been reported from both cross-sectional and longitudinal viewpoints [2-6, 9-14], where *F0* was measured mostly from running speech rather than sustained phonation of a vowel as was employed in this paper, since it is considered that people unfamiliar with acoustic phonetics can speak more naturally in running speech. The two *F0* measures, however, are expected to be correlated with each other. Figure 3 illustrates the relationship between the fundamental frequencies measured from sustained vowel and running speech, “my name is ***,” of the same speakers, where average *F0* was measured over a vowel [a] segment in the second mora of /name/ “name.” Both *F0* values are clearly correlated and the sustained vowel *F0* seems to be about 8.7 Hz higher than that in running speech. All the females who revealed significant *F0* changes resulted in *F0* falling with aging, supporting most of the reports in the past. For males, however, heterogeneous measurements were reported, with *F0* falling or rising with aging. In our data, 90% of the males showed no significant *F0* change but 10% showed *F0* falling; no *F0* rising was observed.

Although several previous cross-sectional studies reported an increase of jitter with aging [3, 9], it seems stable in our measurements irrespective of aging for both males and females; no males showed a significant change, and only 15.8% of females did so. Four females showed decreasing jitter values but the reason for this tendency is not clear yet. Shimmer and noise energy at higher frequencies, on the other hand, tend to increase consistently with aging. These longitudinal characteristics are related to changes in the vibratory behavior of the vocal folds due to physiological alterations of the vocal folds and respiratory system, but detailed discussions are left for future study.

5. Conclusions

The longitudinal change of selected voice source parameters was evaluated by the statistical t-test, revealing that: 1) strong individuality existed for significant longitudinal changes in the acoustic parameters and a relatively small number of subjects did show significant age-related changes, 2) *F0* falling is a more typical tendency of vocal aging of females than males, while *F0* rising, which has been pointed out in the past for males, was not found, 3) shimmer is a more observable indication of vocal aging than jitter, and 4) glottal noise in a high-frequency region tends to increase with aging.

6. Acknowledgements

This research was in part supported by a Grant-in-Aid for Scientific Research (No. 19500155 and 22500147) from the Ministry of Education, Culture, Sports, Science and Technology.

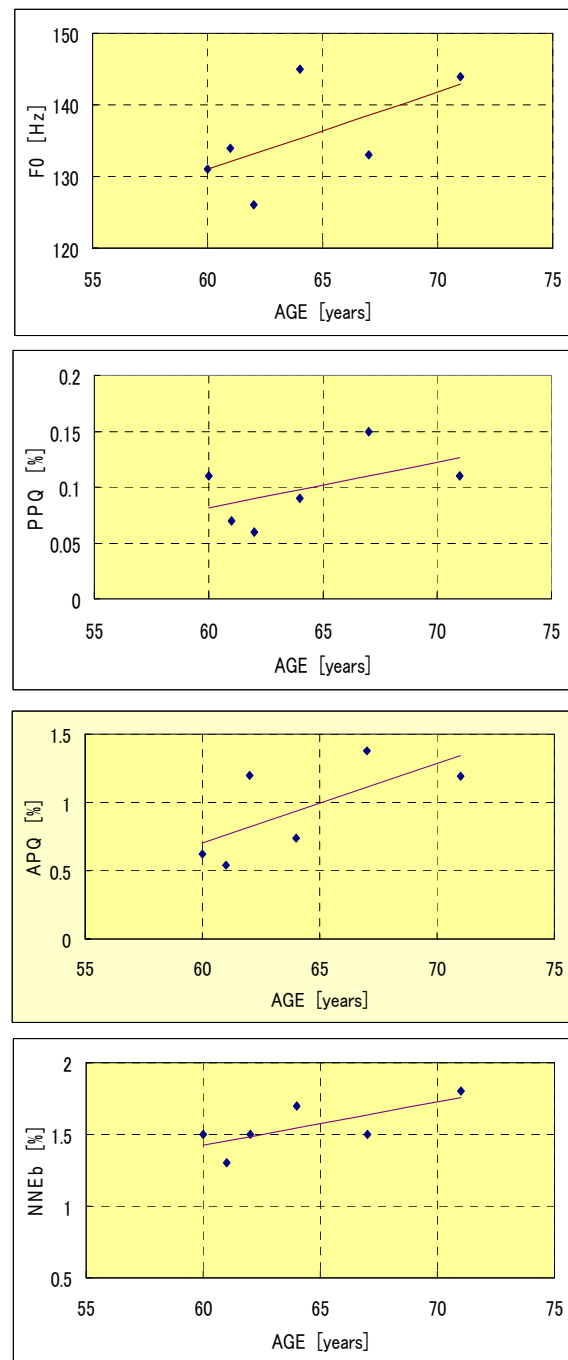


Fig. 2 Same as Fig. 1, for a male subject.

7. References

- [1] Linville, S. E., “Vocal aging,” Singular Thomson Learning; San Diego, 2001.
- [2] Stoicheff, M. L., “Speaking fundamental frequency characteristics of nonsmoking female adults,” *J. Speech Hear. Res.* 24:437–441, 1981.

- [3] Orlikoff, R., "The relationship of age and cardiovascular health to certain acoustic characteristics of male voices," *J. Speech Hear. Res.* 33:450–457, 1990.
- [4] Brown, W., Jr., Morris, R., Hollien, H. and Howel, E., "Speaking fundamental frequency characteristics as a function of age and professional singing," *Journal of Voice*, 5(4):310–315, 1991.
- [5] Linville, S. E., "The sound of senescence," *Journal of Voice*, 10(2):190–200, 1996.
- [6] Decoster, W. and Debryne, F., "The aging voice: changes in fundamental frequency, waveform stability and spectrum," *Acta Otolaryngol, Belgium*, 51(2):105–112, 1997.
- [7] Ferrand, C. T., "Harmonics-to-noise ratio: an index of vocal aging," *Journal of Voice*, 16(4):480–487, 2002.
- [8] Brückl, M. and Sendlmeier W., "Aging female voices: an acoustic perceptive analysis," *Proceedings of VOQUAL'03, Geneva, August 2003.*
- [9] Nishio, M., Tanaka, Y. and Niimi, S., "Analysis of age-related changes in the acoustic characteristics of the voice," *Japan Journal of Logopedics and Phnoatrics*, 50(4):6–13, 2009.
- [10] de Pinto, O. and Hollien, H., "Speaking fundamental frequency characteristics of Australian women: then and now," *Journal of Phonetics*, 10:367–375, 1982.
- [11] Russell, A. and Penny, L., "Speaking fundamental frequency changes over time in women: a longitudinal study," *J. Speech Hear. Res.* 38:101–109, 1995.
- [12] Decoster, W. and Debryne F., "Longitudinal voice changes: facts and interpretation," *Journal of Voice*, 14(2):184–193, 2000.
- [13] Verdonck-de Leeuw, I. M. and Mahieu H. F., "Vocal aging and the impact on daily life: a longitudinal study," *Journal of Voice*, 18(2):193–202, 2004.
- [14] Harrington, J, Palethorpe, S. and Watson C. I., "Age-related changes in fundamental frequency and formant: a longitudinal study of four speakers," *Proc. Interspeech 2007, Antwerp, Belgium, August 2007.*
- [15] Brückl, M., "Women's vocal aging: a longitudinal approach," *Proc. Interspeech 2007, Antwerp, Belgium, August 2007.*
- [16] Kasuya, H., Masubuchi, K., Ebihara, S. and Yoshida, H., "Preliminary experiments on voice screening," *Journal of Phonetics*, 14:463–468, 1986.
- [17] Mori, H., Otawa, H., Ono, T., Ito, Y. and Kasuya, H., "Internet-based acoustic voice evaluation system for screening of laryngeal cancer," *J. Acoust. Soc. Japan*, 62(3):193–198, 2006.
- [18] Koike, Y., "Application of some acoustic measures for the evaluation of laryngeal dysfunction," *Studia Phonologica (Kyoto University)*, VII:17–23, 1973.
- [19] Kasuya, H., Ogawa, S. and Kikuchi, Y., "An adaptive comb filtering method as applied to acoustic analysis of pathological voice," *Proc. Int. Conf. Acoust. Speech Signal Process.*, 669–672, April 1986.

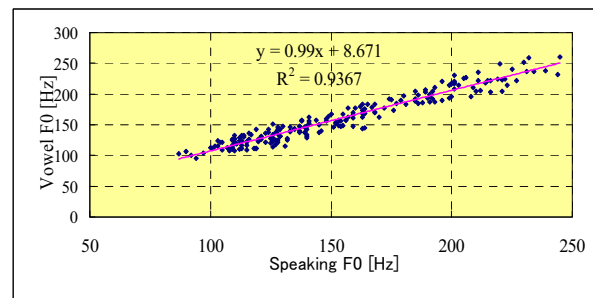


Fig. 3 Relationship between sustained vowel *F0* and *F0* in running speech.

Table 1 Number of subjects who showed a statistically significant change for each of the five acoustic parameters in the 20 males. Numerals in parentheses indicate the number of subjects showing a negative gradient, *i.e.* fall or decrease of the parameter.

Starting age	Number	<i>F0</i>	<i>PPQ</i>	<i>APQ</i>	<i>NNEa</i>	<i>NNEb</i>
40s	3					1
50s	6	1(1)		2		1
60s	8	1(1)		3	1	5
70s	3					
Total	20	2(2)	0	5	1	7

Table 2 Same as Table 1, for the 38 female subjects.

Starting age	Number	<i>F0</i>	<i>PPQ</i>	<i>APQ</i>	<i>NNEa</i>	<i>NNEb</i>
30s	4	2(2)		1		3
40s	10		3(3)	4	2(1)	4
50s	16	6(6)	1	5	1	7
60s	7	2(2)	2(1)	1	2(1)	
70s	1					
Total	38	10(10)	6(4)	11	5(2)	14