Development and Evaluation of Pitch Adjustable Electrolarynx

Yoshinobu Kikuchi & Hideki Kasuya***

*Department of Radiological Science
International University of Health and Welfare, Japan
ykikuchi@iuhw.ac.jp

**Department of Electrical and Electronic Engineering
Utsunomiya University, Japan
kasuya@klab.jp

Abstract

We constructed a prototype of an electrolarynx whereof the fundamental frequency (hereafter pitch for simplicity) can be adjusted by an up-down or left-right finger movement, and carried out some comparison experiments on its operation. As a result, we found that the left-right finger movement could operate the device more simply and in a shorter time. In both cases, in the prototype electrolarynx, pitch adjustment and ON/OFF control of the vibration could be made independently, and utterances could be started and finished with a desired pitch. Hence, we found that an initial-accented pitch pattern and utterances with rising pitch at the end of a phrase such as a question phrase, were easily produced.

1. Introduction

An artificial larynx or electrolarynx has been developed as a means of allowing people who have had their larynx removed to talk. The electrolarynx can be used relatively soon after a surgical operation, and as it is simple and hygienic, and is relatively easy to master, it is widely employed. However, since a large number of such devices on the market do not permit intensity or pitch to be varied, it is desired to develop an electrolarynx which allows a user skilled in its operation to vary the pitch of his/her voice. Research is now in progress on this subject in several institutions with a view to commercialization. We may classify the methods currently available as follows:

1) Automatic pitch pattern generation method
In normal speech, the pitch of the beginning of the speech is high, and the pitch gradually falls off with time. There is a product which incorporates this tendency. A high pitch frequency is produced when the switch of the electrical larynx is turned on, and the pitch then falls [1]. This is more natural than a flat pitch frequency response, but as the pitch variation pattern is fixed, the device cannot be used with the intention of adding various intonation patterns, in particular question form.

Also, although only in the research stage, studies are in progress to produce a pitch pattern within the context of a simple speech synthesis model based on the magnitude of an operating amount [2]. The output of a finger pressure sensor is determined, a phrase command (command at the beginning of an utterance) and accent command are extracted from this magnitude, and a pitch pattern is thus generated. Complex pitch control can be realized by means of a simple operation, but talking or singing outside the limitations of the model is impossible.

2) Direct response method
This is a method of changing a pitch by relating to some operating amount and various operation amounts are proposed. In one product, pitch control is performed directly by the output of an expiration pressure sensor pressed against the tracheal opening [3]. Pitch control will probably become possible as the user becomes accustomed to adjusting the expiration pressure, but both hands are needed to hold the electrolarynx while pressing the expiration pressure sensor correctly in contact with the tracheal opening, so there are problems concerning its ease of operation.

In another product, a pressure sensor is built into a push button, and the pitch is controlled by the force (finger pressure) with which this button is pressed [4]. As this device works by pressure, the displacement amount of the finger for adjusting pitch is zero in principle, but fine control is difficult.

In these products, there is a direct correspondence between operating amount and pitch, so the pitch can be varied as desired with increasing practice. However, as vibration ON/OFF is determined by a threshold value which is set relative to the operating amount, speech cannot be started or ended with a desired pitch. Another study is being performed to establish a correlation between operating amount and pitch using various functions [5], but there seems to be no major significant difference.

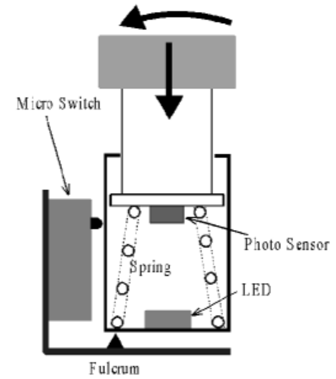
From the above, it is seen that if we wish to adjust pitch as desired, the direct correspondence model is superior, and taking account of songs and question forms, pitch and vibration ON/OFF must be controlled independently from one another [6, 7]. Hereafter, we shall discuss a method to make a finger displacement amount correspond directly to pitch frequency.

2. Pitch control mechanism

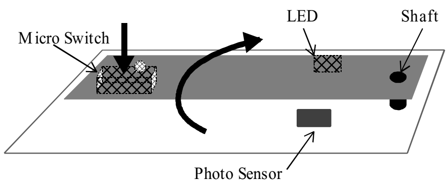
As the thumb can move freely up and down or left and right, we decided to use it as the finger for controlling the pitch of the electrolarynx. Herein, we developed the two mechanisms shown below.

In one of these, the pitch period is controlled by the up-down displacement amount of the finger, and vibration ON/OFF is controlled by the left-right movement. Figure 1a) shows this construction. In the other method, the pitch period is controlled by the left-right displacement amount of the finger, and vibration ON/OFF by the up-down movement. Figure 1b) shows this construction. In Figure 1, the bold arrow shows the direction in which the operating mechanism can move, and it is pressed by a spring in the opposite direction to that of the arrow. As seen from the diagram, the distance between a light-emitting diode (LED) and a photo sensor varies depending on the movement of the finger, so the light amount entering the photo sensor varies. This variation is converted to an electrical signal, so the displacement amount can be detected. This detected electrical signal is used as a pitch control signal. Vibration ON/OFF is performed by a micro switch, and this becomes a voicing command.

As shown in Figure 2, these adjustment mechanisms are connected to a personal computer via an interface, the pitch being controlled by a control program. Figure 3 shows a timing chart in pitch adjustment. The overall system operates in synchronism with a 10 kHz clock. The A/D converter sampling period is also 10 kHz. When voicing commands are OFF, voicing commands are repeatedly searched. This is the state shown by the white circles "O" in Figure 3. When voicing commands are ON, a pitch operating value is input, and the pitch period corresponding to this value is calculated. This is the state shown by the black circles in Figure 3. In this test, the vibrating body was a hammer striking a vibration plate.



a) *Method by up-down displacement of the finger.*



b) *Method by left-right displacement of the finger.*

Figure 1: *Pitch adjustment mechanisms which have two degrees of freedom.*

To obtain a fixed vibration intensity, the time of applying a voltage to a drive coil was fixed, and the pitch period was adjusted by the OFF time. In this way, the vibrating body was controlled by a PFM (Pulse Frequency Modulation) method where the pitch period was varied by adjusting the OFF time of the drive coil. In the test system, the OFF time was measured by counting the 10 kHz clock, and the pitch adjustment precision was 100 microseconds. Due to this, the frequency adjustment precision was 1 Hz in the vicinity of 100 Hz, and 9 Hz in the vicinity of 300 Hz.

3. Experiments

To investigate the ease of operation of the prototype electrolarynx, a learning test was performed by healthy subjects. The aim of the test was to copy the pitch variation of various utterances which were made previously. In the test, it was not compulsory to place the electrolarynx in contact with the throat. Measurements were performed after 30 minutes practice.

The subjects were five female students intending to become speech therapists, and although they had knowledge of speech production, they were not familiar with the operation of the electrolarynx. The sentences uttered were conversation sentences used in daily life. Question sentences were used to investigate the features of this electrolarynx, where utterances could be started or finished with a desired pitch. Table 1 shows these sentences. Sentence 1 is a sentence without pauses, and in the operation of the electrolarynx, adjustment was continuous. As this is a question sentence, the pitch of the end phrase is high. Sentence 2 consists of two sentences with a pause. In the latter half of Sentence 2, voiced and silent explosive sounds are continuous.

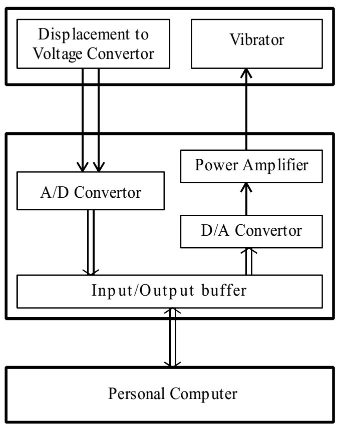


Figure 2: *Block diagram of a pilot electrolarynx system.*

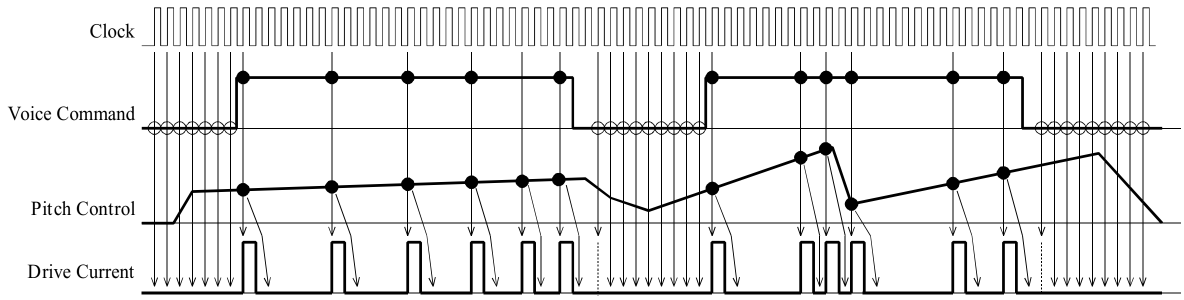


Figure 3: Timing chart of pitch control.

Table 1: Sentences used in the experiments.

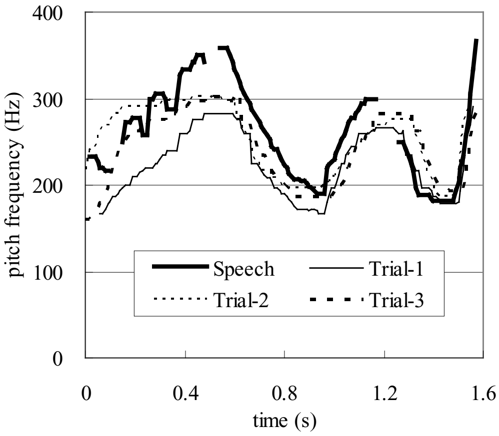
Sentence 1	Ohirugohanwa dohsuru? (What about lunch?)
Sentence 2	Ohayou gozaimasu. Odekake desuka? (Good morning, are you going out?)

4. Results and Discussion

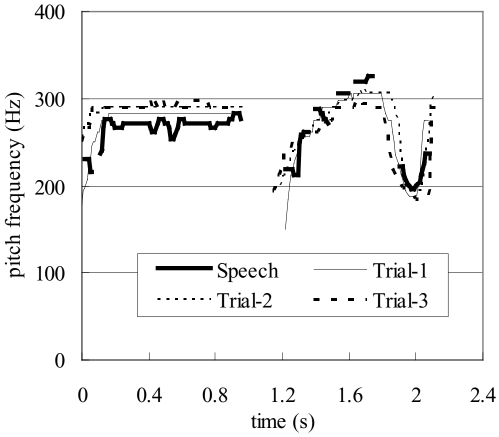
The pitch variation was examined based on speech utterance and the vibration sounds obtained by operating the electrolarynx. Figures 4 and 5 show examples of the pitch pattern obtained. The bold line is the pitch pattern of natural conversational sentences, and the other lines are pitch patterns obtained by operating the electrolarynx. The results are obtained for control in the up-down direction of Figure 1a).

Figure 4 shows the results for Subject 1. a) is the pitch pattern of Sentence 1. In the figure, the time axis and frequency axis are not adjusted. Normally, in parts where there are no voiced sounds, the electrolarynx does not stop vibrating and a continuous pitch variation is observed. The pitch range of an utterance is 180-370 Hz, which is substantially identical to the adjustment range of the electrolarynx. Figure 4b) shows the results for Sentence 2. As the length of the pauses is different, the pause length alone is adjusted by visual observation so that overlap is improved. Figure 5 shows the results for Subject 2. a) shows the results for Sentence 1. Whereas the pitch of utterances is 200-380 Hz, the pitch of the electrolarynx is shifted to the lower end. For Subject 3, similar results are obtained. As can be seen from the results for these two subjects, the absolute magnitude and variation range of the pitch are different from those of utterances, but similar pitch variations could be produced. From the tests performed on these five subjects, four subjects concluded that the left-right pitch control mechanism of Fig. 1b) was easier to use.

As a previous research reported that an on/off control of the vibration for consonant parts of an utterance improved naturalness of the utterance[5], we tested if simultaneous control of on/off of the vibration and pitch change could be attained with the proposed left-right pitch control mechanism. Figure 6 shows results for utterance "odekakedesuka ?" of Sentence 2. It is seen that although the pitch adjustment operation is somewhat sluggish, the control went well.

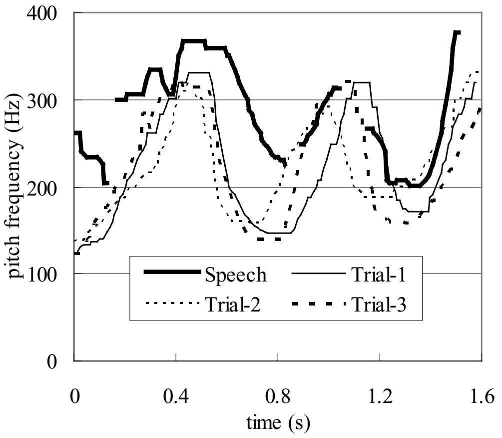


a) *Results of sentence 1*

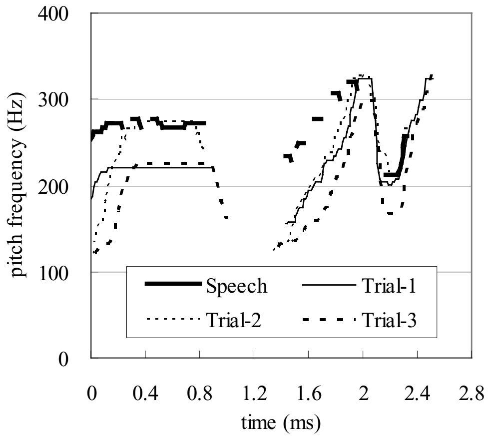


b) *Results of sentence 2*

Figure 4: *The analysis results of Subject 1.*



a) *Results of sentence 1*



b) Results of sentence 2

Figure 5: *The analysis results of Subject 2.*

5. Conclusion

We have described how two prototype pitch adjustment mechanisms were used in an electrolarynx, one as voiced commands and the other as a pitch control, where pitch control was performed by finger movement. It was found that operations relating to pitch control were learnt in a short time of about 30 minutes. By making pitch adjustment independent from vibration ON/OFF commands, the pitch of the last phrase could easily be made to rise, as in a question sentence. It was also found that four out of five subjects considered pitch adjustments in the left-right direction were easier to perform.

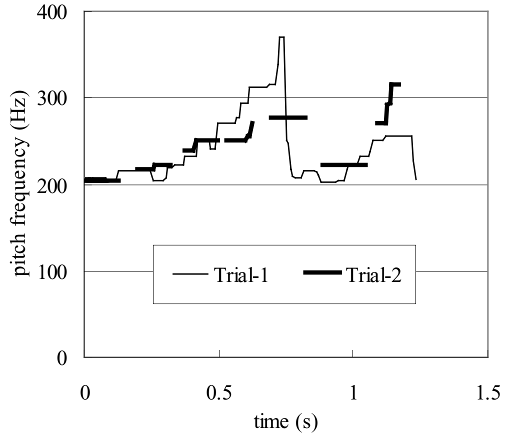


Figure 6: *Results for the continuous explosive sound parts. Trial-1 shows the result when voiced command is continuous and Trial-2 is the result when it was interrupted.*

6. References

[1] SECOM homecare system company. The treatment manual of MYVOICE. (http://medical.secom.co.jp/myvoice/)

[2] Takahashi, H.; Nakao, M.; Ohkusa, T.; Hatamura, Y.; Kikuchi, Y.; Kaga, K., 2001. Pitch control with finger pressure for electrolaryngeal or intra-mouth vibrating speech. *Jp. J. Logopedics and Phoniatrics,* 42(1), 1-8.

[3] Uemi, N.; Ifukube, T.; Takahashi, M.; Matsushima, J., 1995. Development of an electrolarynx having a pitch frequency control function by using expiration pressure. *Jp. Soc. Medical and Biological Engineering.* 33(1), 7-14.

[4] Griffin laboratories., TruTone users guide. (http://www.griffinlab.com/)

[5] Kakita, Y.; Hirama, J., 1989. Controles of prosodic information and voiceless consonants for the electronic larynx. *Technical report of IEICE*. SP88-148, 25-30.

[6] Kikuchi, Y.; Kasuya, H., 2002. Development of electrolarynx which has F0 control feature. Proc. *Autumn Meeting Acoust. Soc. Japan*. 2-10-5, 295-296.

[7] Kikuchi, Y.; Kasuya, H., 2002. Consideration of F0 control of electro larynx. *Technical report of IEICE*. SP2001-106, WIT2002-47.