A Comparative Study of Glottal Data from Normal Adults Using Two Laryngographs[1)]

Byunggon Yang[*], Soo-Geun Wang[**], Soon-Bok Kwon[**]
[*]Dongeui University, [**]Pusan National University

ABSTRACT

A laryngograph was developed to measure the open and closed movements of vocal folds in our laboratory. This study attempted to evaluate its performance by comparing its glottal data with that of the original laryngograph. Ten normal Korean adults participated in the experiment. Each subject produced a sustained vowel /a/ for about five seconds. This study compared f0 values, contact quotients of the duration of closed vocal folds over one glottal pulse, and area quotients of the closed over open vocal folds derived from glottal waves using both the original and new laryngographs. Results showed that the mean and standard deviation of the two laryngographs were almost comparable with a correlation coefficient 0.662 but minor systematic shift below those of the original laryngograph was observed. The absolute mean difference converged into 1 Hz, which indicates a possibility of adopting some threshold of rejecting inappropriate pitch values beyond a threshold value. The contact quotient of the normal subjects came out slightly over the 50% in a citation speech. Finally, the area quotient converged into 1. We will pursue further studies on the abnormal patients in the future.

Key words: laryngograph, glottal waveform, pitch variation

1. Introduction

The source of speech output can be best traced by using a laryngograph since it directly reflects the vocal fold movements through the electrical impedance between two electrodes placed on the superficial layer of the thyroid cartilage (Fourcin 1981). When the two vocal folds are contacted, a strong current will pass while being blocked by the intervening air when they are open (Hayward 2000). It is useful to measure the pitch values against background noises because it collects only the electrical currents across the vocal folds. As a result, we may minimize pitch detection errors which come from the voice waveform mixed with background noises. A therapist can help clients practice voice exercises by using information of

1) This study was supported by a grant of the Korea Health 21 R&D Project, Ministry of Health & Welfare, Republic of Korea (02-PJ1-PG10-31401-0005).

vocal fold activities on a graphic display.

Recently, we have developed a laryngograph in our research laboratory (Song 2003). This study attempts to evaluate the performance of the new laryngograph by comparing its glottal data with those of the original device. We will examine the laryngographic data collected from ten normal Korean males using the two laryngographs.

Normally only one laryngograph can collect direct input from the two electrodes attached to the thyroid cartilages of a subject. Thus, comparable data was collected by the original laryngograph and the new one in sequence. In order to minimize the acoustic variation within a speaker, we decided to control the production of each subject by asking him to listen to a model synthesized vowel /a/ and to repeat it at the same pitch level. The comparison may raise questions and problems, which will lead to further studies on this topic in the future. In addition, this study may provide basic data on the glottal waveform of normal adults.

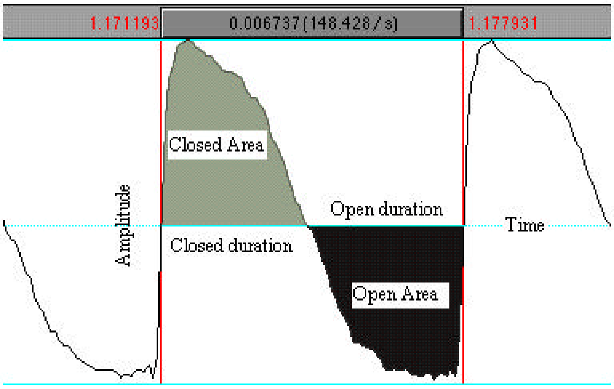
2. Method

Ten subjects participated in this experiment. They were graduate students of the Pusan National University Hospital. None of them reported any speaking or hearing problems. Their age ranged from 26 to 33 years old (average=29.6;s.d.=2.63).

Each subject produced sustained vowel /a/ (greater than 5 seconds) a few seconds after listening to the stimulus in a quiet office. To elicit somewhat similar pitch level, a token of synthesized /a/ with an f0 of 130 Hz and the duration of 10 seconds was presented via a headphone. The stimulus was synthesized using SenSynPPC 1.0 (Yang 1995). The amplitude and formant values were obtained from one of the author's production and the other input values were set as default ones. The recording was made on a personal computer at a sampling rate of 11,025 Hz per second. In order to examine individual variation, five productions per each subject were stored on the computer.

A total of 40 tokens (4 productions x 10 subjects) were selected for analysis using Praat, a speech analysis software. We made three kinds of measurements: f0 values per token, contact quotients, and area quotients. Figure 1 illustrates a laryngographic waveform.

Figure 1. The first subject's glottal waveform showing two pitch pulses. X-axis denotes time while y-axis does amplitude. Upper numbers indicate sector duration in seconds. 148.428 is f0 value of the selected sector in Hz. The closed area is filled in gray while the open area in black.



The contact quotient was determined from the closed duration to the open duration in time. In order to compare subtle wave variations, an area quotient was also calculated from the closed area over the open area (Dejonckere and Lebacq, 1985). They made an electroglottographic study of normal adults and patients with vocal nodules and reported the area quotient of the glottal waveform of the patients with vocal nodules significantly reduced compared to that of the normal subjects.

Those values were collected using a Praat script (Appendix). The script read the wave file and bandpassed it within the range of 50~4,000 Hz to remove high frequency noises. Then, the laryngographic data were traced to find the starting and ending point of each glottal pulse by searching the crossing point at which the positive or negative sign of amplitude value changes into the opposite sign. The current script did not include the first derivative of the glottal waveform which is a good indicator of the timing of laryngeal opening and closure (Sercarz et al. 1992). Then, the pitch values were determined from the duration of each glottal pulse. Since the duration of each production was not the same, a decision to examine the first 250 glottal pulses per token was made to compare two different measurements. Four out of five productions were adopted to derive statistical results in order to avoid obvious measurement errors, too high or too low pitch values within each subject. For example, a subject showed an average of 333 Hz for his f0 values in one production but that of the other four productions was stable around 154 Hz. The reason might be caused by the loose electrodes to the neck skin while producing the vowel. Another subject had an average pitch value of 121 Hz for the first token while that of the other four productions amounted to 145 Hz. Probably the subject could not tune to the stimulus at the beginning.

In order to test the validity of pitch values from the script we compared the first subject's f0 values either by the Praat pitch analysis every 0.0067 ms which almost matched each glottal pulse or by the script. The average difference between both methods came out as -2.02 Hz within the first 400 points. Notably, there was a normal distribution around the average difference. 50% (197 points) of the 400 points showed a difference of -2 Hz while 21% (84 points) were placed 1 Hz below or above the mean. Then, 5% followed 2 Hz below or above the mean. The correlation coefficient between the pitch values by the two methods was 0.735. The pitch shift might be related to different interpolation techniques deriving the pitch values.

3. Results and Discussion

3.1 Comparison of f0 Statistics

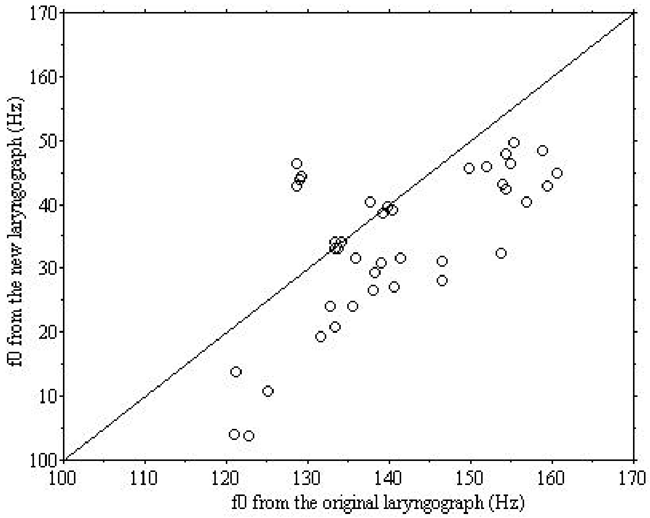
Table 1 lists means and standard deviations of 250 pitch values per token produced by ten Korean adults using the two laryngographs.

Table 1. Means and standard deviations of f0 values collected from ten Korean subjects using both the original and new laryngographs. The units are in Hz.

Subject	f01	sd1	f02	sd2	f03	sd3	f04	sd4
L1	129.2	1.1	128.6	1.1	128.6	1.0	129.0	1.0
L2	156.8	1.9	159.4	1.4	158.8	1.2	160.5	1.1
L3	153.8	1.2	154.3	1.3	155.0	1.5	155.2	1.3
L4	153.9	1.4	154.3	2.0	151.9	2.1	149.9	1.5
L5	140.3	0.7	137.6	2.4	139.3	1.1	139.9	1.0
L6	122.8	1.2	121.0	0.6	125.0	1.3	121.2	0.7
L7	135.9	1.0	139.1	1.1	138.2	1.4	141.4	1.0
L8	146.4	1.1	146.4	1.7	140.5	1.0	138.1	1.4
L9	133.4	0.6	133.3	0.5	134.1	0.8	133.7	0.5
L10	132.8	0.9	135.5	1.4	133.4	1.1	131.6	1.6
Average	***140.5***	***1.1***	***140.9***	***1.4***	***140.5***	***1.2***	***140.0***	***1.1***
N1	144.5	1.3	142.9	4.5	146.4	1.1	143.9	1.2
N2	140.5	0.8	143.0	1.3	148.5	1.0	145.0	1.5
N3	132.3	1.1	142.5	1.1	146.5	1.2	149.7	1.2
N4	143.1	1.8	148.0	1.8	146.0	1.3	145.7	2.1
N5	139.2	1.5	140.4	1.0	138.6	1.7	139.6	1.3
N6	103.7	0.9	104.1	1.8	110.7	1.2	113.7	1.1
N7	131.7	1.5	130.8	1.2	129.4	1.3	131.7	1.1
N8	128.2	1.4	131.0	0.9	127.2	1.3	126.7	1.4
N9	133.1	0.7	134.2	0.9	134.2	0.9	133.0	0.7
N10	124.2	1.3	124.2	1.2	120.7	0.7	119.2	1.0
Average	***132.1***	***1.2***	***134.1***	***1.6***	***134.8***	***1.2***	***134.8***	***1.2***

The average f0 values on the 12th row of the ten subjects by the original laryngograph are almost the same as those on the 23rd row by the new laryngograph. The glottal data from the new device yielded quite similar pitch values to those of the original one. The standard deviation of each token is almost comparable around 1 Hz. L2 has the highest pitch value of 160.5 Hz while N6 lists the lowest pitch of 103.7 Hz. N1 lists the highest standard deviation (4.5 Hz) while L9 shows the lowest deviation (0.5 Hz). To examine the pitch values of the two devices in detail, we plot each mean f0 value of the four productions altogether in Figure 2.

Figure 2. f0 values from the original laryngograph and the new laryngograph.



In Figure 2, one can notice that the pitch values are spread around the line of equity. Generally, those values from the laryngograph are lower than those of the new device. The correlation coefficient is 0.662. There are a few points which exactly match but we can note some systematic down shift from the values of the new laryngograph. Since we could not measure the values at the same time on the same subject, we could expect some variation within each subject. Each individual may control pitch values if he repeats the synthetic stimulus over a headphone. However, since each individual has a different perceptual range of the same pitch, the pitch values are supposed to vary slightly on an individual basis. Yang (2001) reported that twelve male and female subjects perceived the same sound quality within the range of 10.5 Hz of pitch variation. Further studies on trained singers may be interesting to see how exactly they can mimic various pitch values over a headphone.

Table 2. lists means of absolute difference between adjacent pitch values.

Table 2. Comparison of the mean absolute difference between adjacent pitch values of four tokens per subject. The units are in Hz.

Subject	Diff1	Diff2	Diff3	Diff4
L1	0.82	1.00	0.93	0.88
L2	1.64	1.20	1.29	1.00
L3	1.31	1.07	1.17	1.18
L4	1.10	1.24	1.20	1.12
L5	0.65	0.96	0.99	0.73
L6	0.71	0.52	0.83	0.54
L7	0.98	1.02	1.00	0.80
L8	1.05	1.16	0.80	1.07
L9	0.60	0.57	0.66	0.59
L10	0.79	1.08	0.78	0.79
Average	***0.96***	***0.98***	***0.97***	***0.87***
N1	1.17	0.92	1.24	1.10
N2	0.68	0.98	0.95	1.29
N3	0.90	1.00	1.01	1.21
N4	0.99	1.12	1.12	1.34
N5	1.04	0.79	0.96	0.87
N6	0.73	1.04	0.73	1.17
N7	1.08	0.94	1.01	0.88
N8	0.90	0.69	0.75	0.90
N9	0.70	0.82	0.74	0.65
N10	0.79	0.81	0.56	0.56
Average	***0.90***	***0.91***	***0.91***	***1.00***

In Table 2, an absolute mean difference of around 1 Hz occurs between the adjacent pitch values. The difference may be comparable to jitter or pitch variation (Wang et al. 1999). The highest difference came out as 1.64 Hz (L2) while the lowest difference did 0.52 Hz (L6). The pitch values of a normal subject did not change greatly. This result may be applicable to the pitch-sifting algorithm in which the pitch values higher than a certain threshold may be rejected as obvious measurement errors. Also, those pitch variations may be applicable to the synthesis of natural speech output. The same f0 parameter across the whole duration of a vowel may sound like a mechanic speech. Next, we will examine the contact quotient of each production.

3.2 Comparison of Contact quotients

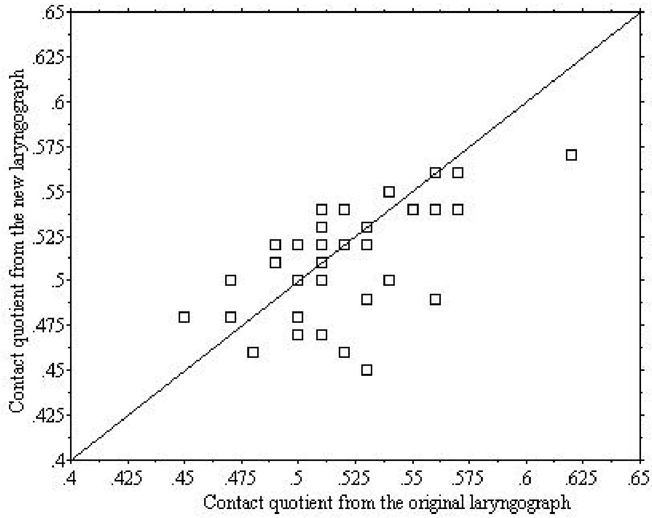
Table 3 summarizes the contact quotients of f0 values produced by the subjects using the two laryngographs.

Table 3. Means and standard deviations of contact quotients of f0 values produced by ten subjects.

Subject	CQ1	CQsd1	CQ2	CQsd2	CQ3	CQsd3	CQ4	CQsd4
L1	0.51	0.009	0.53	0.007	0.51	0.006	0.51	0.009
L2	0.50	0.011	0.53	0.008	0.54	0.009	0.54	0.007
L3	0.55	0.010	0.56	0.009	0.56	0.007	0.62	0.008
L4	0.50	0.005	0.50	0.007	0.49	0.006	0.49	0.006
L5	0.48	0.009	0.52	0.018	0.51	0.006	0.53	0.008
L6	0.49	0.007	0.45	0.006	0.50	0.008	0.47	0.007
L7	0.56	0.007	0.57	0.008	0.55	0.008	0.57	0.009
L8	0.49	0.007	0.47	0.015	0.56	0.011	0.47	0.017
L9	0.52	0.005	0.51	0.005	0.52	0.005	0.51	0.005
L10	0.52	0.006	0.51	0.007	0.53	0.013	0.51	0.011
Average	***0.51***	***0.008***	***0.51***	***0.009***	***0.53***	***0.008***	***0.52***	***0.009***
N1	0.50	0.009	0.53	0.013	0.51	0.011	0.53	0.010
N2	0.47	0.006	0.49	0.007	0.50	0.008	0.55	0.027
N3	0.54	0.007	0.54	0.007	0.56	0.010	0.57	0.007
N4	0.50	0.005	0.52	0.005	0.52	0.007	0.51	0.006
N5	0.46	0.012	0.46	0.009	0.47	0.014	0.45	0.008
N6	0.51	0.007	0.48	0.011	0.48	0.008	0.48	0.014
N7	0.54	0.012	0.54	0.013	0.54	0.019	0.56	0.011
N8	0.51	0.005	0.50	0.005	0.49	0.008	0.50	0.008
N9	0.54	0.007	0.52	0.006	0.54	0.013	0.54	0.007
N10	0.52	0.008	0.54	0.006	0.52	0.008	0.53	0.006
Average	***0.51***	***0.008***	***0.51***	***0.008***	***0.51***	***0.010***	***0.52***	***0.010***

From Table 3, one can observe that there is not much difference in the average quotient of the two equipments. On an average, the contact quotient of the duration of contacted vocal folds to the total duration of the glottal pulse amounted to just a little over the half. That means that the closed duration is slightly longer than the open duration. In detail, 0.62 of L3 lists the highest quotient while 0.45 of L6 does the lowest. The standard deviation ranges from 0.005 to 0.027. These normal subjects produced almost comparable contact quotients from the two devices. Here again, we plot the individual values from the two laryngographs in Figure 3.

Figure 3. Contact quotients obtained from the original laryngograph against those of the new laryngograph.



The values were spread around the line of equity. The correlation coefficient was 0.576 which shows somewhat weaker correlation between the two data. Generally, the data cluster around 0.5 which is within the range of the modal voice between 30 and 55% (Carlson and Miller 1998; Yang 1996).

3.3 Comparison of Area Quotients

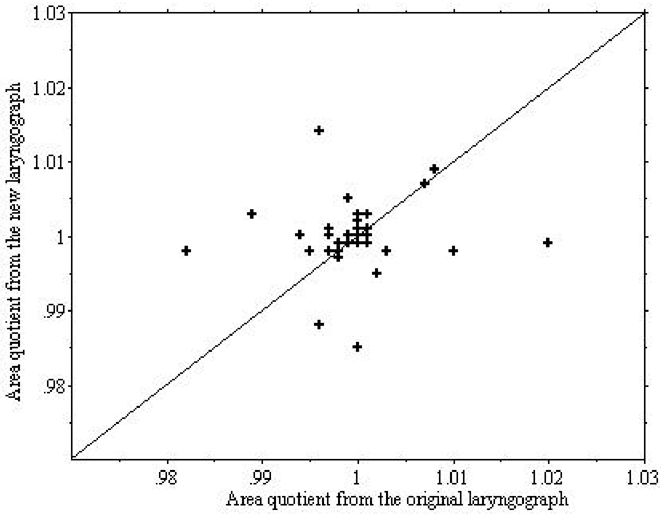
Table 4 shows the means and standard deviations of area quotients of f0 values from the subjects. If we compare the area quotients, we can make more subtle comparisons of the laryngographic waveforms.

Table 3. Means and standard deviations of the area quotients of f0 values obtained from ten subjects using the two laryngographs.

Subject	AR1	ARsd1	AR2	ARsd2	AR3	ARsd3	AR4	ARsd4
L1	1,001	0,032	1,002	0,027	1,000	0,026	1,001	0,041
L2	0,982	0,011	0,995	0,017	0,994	0,012	0,989	0,008
L3	1,007	0,081	1,001	0,014	1,000	0,016	1,001	0,010
L4	1,000	0,006	0,994	0,020	1,000	0,006	1,000	0,007
L5	1,001	0,016	0,998	0,012	0,998	0,011	0,998	0,011
L6	0,997	0,008	0,997	0,009	0,997	0,018	0,997	0,008
L7	0,996	0,028	1,000	0,020	1,008	0,029	0,996	0,068
L8	0,997	0,070	1,010	0,189	1,003	0,080	1,020	0,167
L9	0,998	0,008	0,999	0,008	0,999	0,008	0,999	0,008
L0	1,000	0,007	1,000	0,006	1,000	0,007	0,999	0,007
Average	***0.998***	***0.027***	***1.000***	***0.032***	***1.000***	***0.021***	***1.000***	***0.034***
N1	1,000	0,045	0,995	0,070	1,003	0,047	1,001	0,034
N2	0,998	0,019	0,998	0,015	1,000	0,046	1,003	0,103
N3	1,007	0,097	1,003	0,069	1,000	0,013	1,003	0,036
N4	1,001	0,007	1,000	0,009	1,002	0,051	1,001	0,020
N5	0,999	0,051	0,998	0,027	0,997	0,015	0,998	0,038
N6	1,001	0,025	1,000	0,015	0,998	0,026	1,000	0,021
N7	1,014	0,277	0,985	0,077	1,009	0,128	0,988	0,044
N8	0,998	0,005	0,998	0,005	0,998	0,016	0,999	0,018
N9	0,999	0,016	0,999	0,019	1,005	0,129	1,000	0,042
N0	1,000	0,011	0,999	0,008	1,000	0,005	0,999	0,004
Average	***1.002***	***0.055***	***0.998***	***0.031***	***1.001***	***0.048***	***0.999***	***0.036***

Generally, the data from the two devices showed almost comparable. On an average the values converged into 1 while the standard deviation ranges from 0,004 (N10) to 0,277 (N7). The value 1 on the table means the same area value of the open and closed vocal folds. If it is higher than 1, it will mean a wider area of the closed vocal folds. Figure 4 plots each individual area quotient.

Figure 4. Plot of area quotients collected from the original laryngograph and the new laryngograph.



It seems that almost no correlation between the two data exists but we find that they are closely clustered together. Therefore, we may be careful not to conclude that the two laryngographs worked differently. The same applies to the comparison of the contact quotients and the area quotients to find an almost no correlation because those values were clustered closely. If we had included abnormal subjects, we might have obtained somewhat stronger correlation.

4. Conclusion

This study compared the f0 values, the contact quotients of the duration of closed vocal folds over one glottal pulse, and the area quotients of the closed over open vocal folds derived from glottal waves produced by ten healthy Korean adults using both the original laryngograph and a new device. Results showed that the mean and standard deviation of the two laryngographs were almost comparable with a correlation coefficient 0.662 but some systematic down shift from the original laryngograph was observed. The absolute mean difference converged into 1 Hz which indicates a possibility of adopting some threshold of rejecting inappropriate pitch values beyond a threshold value. The contact quotient of the normal subjects came out slightly over the 50%. The area quotient converged into 1. There was almost no correlation between those values from the two laryngographs in terms of contact or area quotients, which might be attributed to the clustered data. Further studies on the abnormal patients will be desirable in the future.

References

Carson, E. and D. Miller. 1998. "Aspects of voice quality: display, measurement and therapy." *International Journal of Language and Communication Disorders* Vol. 33, Supplement. 304-309.

Dejonckere, P.H. and J. Lebacq. 1985. "Electroglottography and vocal nodules: an attempt to quantify the shape of the signal." *Folia phoniat.* 37, 195-200.

Fourcin, A. J. 1981. "Laryngographic assessment of phonatory function in C." L. Ludlow and M. O. Hart (eds.), *Proceedings of the Conference on the Assessment of Vocal Pathology, ASHA Rep.* No. 11, 116-127. Washington D. C.: American Speech and Hearing Association.

Hayward, K. 2000. *Experimental Phonetics.* Harlow, England: Pearson Education.

Jo, C., Wang, S., Yang, B. 1999. "A study on the diagnosis of laryngeal diseases by acoustic signal analysis." *Korean Journal of Speech Sciences,* Vol. 5, 151-65.

Song, I. 2003. *Differential diagnosis of laryngeal cancer using electroglottograph v1.0.* Ph. D. Thesis, Pusan National University.

Wang, S., Baek, M., Yang, B., Jo, C., Park, H., Kweon, S. 1999. "Acoustic parameters for the early detection and differential diagnosis of pathologic voice." *Korean Journal of Otolaryngology,* Vol. 42, 1561-7.

Yang, B. 1995. "A perceptual study of synthesized Korean monophthongs." *Korean Journal of Linguistics,* 20(3), 127-146.

Yang, B. 1996. "An analysis of Korean glottal waves by the laryngograph and a perceptual study of synthesized vowels." *Korean Journal of Linguistics,* 21(4), 1025-1040.

Yang, B. 2001. "Perceptual experiment on number production for speaker identification." *Korean Journal of Speech Sciences,* Vol. 8(1), 7-19.

Appendix: *Praat script*

```
clearinfo
number=1
filename$="lar"+"'number'"
filenamefull$="lar"+"'number'"+".wav"
Read from file... C:\MyDocuments\NormalData\'filenamefull$'
Filter (pass Hann band)... 50 4000 100
filenameband$="'filename$'"+"_band"
select Sound 'filenameband$'
Rename... 'filename$'
   dur=Get duration
   numsample=40000
   samper=Get sample period
   finish='dur'-0.04
   call setting
   call startingpoint
   call endingpoint
   call sylpoint
procedure setting
   i=1
   j=1
   p=1
   sumcontact=0
   sumopen=0
endproc
procedure startingpoint
   valindex=Get value at index... 'i'
     if 'valindex'<0
       while 'valindex'<0 and 'i'<'numsample'
        i='i'+1
        valindex=Get value at index... 'i'
       endwhile
     else
       while 'valindex'>=0 and 'i'<'numsample'
        i='i'+1
        valindex=Get value at index... 'i'
       endwhile
       while 'valindex'<0 and 'i'<'numsample'
        i='i'+1
        valindex=Get value at index... 'i'
       endwhile
     endif
endproc
procedure endingpoint
   j='numsample'
   valindex=Get value at index... 'j'
     if 'valindex'<0
       while 'valindex'<0
        j='j'-1
        valindex=Get value at index... 'j'
       endwhile
       while 'valindex'>=0
        j='j'-1
        valindex=Get value at index... 'j'
       endwhile
```

```
      else
        while 'valindex'>=0
          j='j'-1
          valindex=Get value at index... 'j'
        endwhile
      endif
   endindex='j'+1
endproc
procedure sylpoint
   while 'valindex'<0 and 'i'<'endindex'
      i='i'+1
      valindex=Get value at index... 'i'
   endwhile
   while 'i'<'endindex'
      contactonset'p'='i'
      while 'valindex'>0 and 'i'<'endindex'
        sumcontact='sumcontact'+'valindex'
        i='i'+1
        valindex=Get value at index... 'i'
      endwhile
      cindex='i'-1
      contactoffset'p'='cindex'
      cvalindex=Get value at index... 'cindex'
      openonset'p'='i'
      while 'valindex'<=0 and 'i'<'endindex'
        sumopen=abs('valindex')+'sumopen'
        i='i'+1
        valindex=Get value at index... 'i'
      endwhile
      oindex='i'-1
      openoffset'p'='oindex'
      ovalindex=Get value at index... 'oindex'
      call analyze
      p='p'+1
      sumcontact=0
      sumopen=0
   endwhile
endproc
procedure analyze
   copy contactonset'p' contactonsetpnt
   copy contactoffset'p' contactoffsetpnt
   copy openonset'p' openonsetpnt
   copy openoffset'p' openoffsetpnt
   pulsedur='openoffsetpnt'-'contactonsetpnt'
   f0=1/('samper'*'pulsedur')
   opendur='openoffsetpnt'-'openonsetpnt'
   openratio=('opendur'/'pulsedur')
   coratio='sumcontact'/'sumopen'
   fileappend C:\MyDocuments\NormalData\result
   print 'p''tab$''f0:0''tab$''openratio:3''tab$''coratio:3''newline$'
endproc
```

Received: January 15, 2002.

Accepted: February 28, 2002

▲ Byunggon Yang
English Department, Dongeui University
24 Kayadong, Pusanjingu, Pusan, 614-714, Korea
Tel:+82-51-890-1227, FAX: +82-51-890-1222
e-mail: bgyang@dongeui.ac.kr

▲ Soo-Geun Wang
Department of Otolaryngology, Pusan National University
1-10 Ami-dong, Seo-gu, Pusan 602-739, Korea
Tel : +82-51-240-7331
E-mail : wangsg@pusan.ac.kr

▲ Soon-Bok Kwon
Department of Interdisciplinary program in Biomedical Engineering,
Pusan National University 1-10 Ami-dong, Seo-gu,
Pusan 602-739, Korea
Tel : +82-51-240-7543 H/P : 011-9519-9802
E-mail : sbkwon@intizen.com