

## Glottal Parameters Contributing to the Perception of Loud Voices

Sopae Yi\* · One Good Lee\*\* · Hyung Soon Kim\*\*\*

### ABSTRACT

This paper focused on glottal parameters contributing to the perception of loud voices because energy of a voice is not the only effective factor. We used a formant synthesizer to synthesize loud voices. We divided F0 tilt (the tilt of F0 contour), SQ (Speed Quotient), OQ (Open Quotient) and TL (spectral Tilt Level) into three levels to get different combinations with default values for the other synthesizer parameters. Analysis of listening tests indicated that F0 tilt, SQ, OQ and TL in descending order had significant influence on the perception of loud voices. F0 tilt had a far more significant effect than the others. The influence of SQ increased greatly with the exclusion of F0 tilt as a factor. The interaction between parameters was not significant.

**Keywords :** loudness, LF parameters

### 1. Introduction

Speech has a special value as an interface between humans and machines. Synthesized voices with a phone directory service help people find phone numbers. Synthesized voices also read e-mails through a telephone channel. Many examples of synthesized speech use can be found in everyday life.

Great improvements in modern speech synthesis technology have increased the need for the naturalness of synthesized speech. Even reflecting emotions in synthesized speech has become a research issue. Machine speech using the dynamic characteristics of a natural human voice can also contribute to mitigating listening fatigue.

Since human emotion can be more easily expressed in loud or soft voices than in a normal one, progress on loud voices can contribute to the enhancement of naturalness of a synthetic voice. However, increasing the energy of a voice is not the only effective factor in improving the perception of loud voices. Holmberg et al. measured Speed Quotient (SQ), Open Quotient (OQ), Spectral Tilt Level (TL) and Fundamental Frequency (F0) in natural loud voices. By comparing loud voices with soft voices, the study concluded that F0 and

---

\* Dept. of Cognitive Science, Pusan National University

\*\* Dept. of English Literature and Language, Pusan National University

\*\*\* Dept. of Electronics Engineering, Pusan National University

SQ values increase while OQ and TL values decrease in loud voices [1].

Much research on loud phonation can be found but it is hard to find studies concerning the perception of loud phonation. Therefore, this paper focuses on the perception of loud voices. We used a Klatt88 [2][3] formant synthesizer to produce voices at different volume levels. We also used the LF model as the source signal input. We used the analogue scale method for listeners to quantify the loudness of each sound. We estimated the quantity of the contribution of the glottal parameters (F0, SQ, OQ, TL) contributing to the perception of a loud voice. Loud voices in natural sound tend to have a shorter duration [1]. This paper, however, does not take this 'shortness of duration' phenomena into consideration.

We discuss acoustic characteristics of loud voices in section 2, our experimentation and analysis in sections 3 and 4, and the results in section 5.

## 2. Acoustic Issues of Loud Voices

According to the analysis of natural voices, loud voices have higher F0 (Fundamental Frequency) and SQ and lower OQ than normal voices [1][4]. Figure 1 shows the F0 contours of normal voices and loud voices of [a] vowels. These F0 contours are from two of the ten participants whose F0 values revealed similar patterns. As can be seen in figure 1, the overall F0 values of loud voices are higher and more dynamic than those of the normal counterparts. F0 contours of loud voices rise more steeply in the beginning and fall more steeply in the end than the contours of normal voices.

Figure 2 shows a glottal flow signal, its derivative and the LF parameters.  $U_0$  is the maximum of  $U_g$ .  $E_e$  is the absolute value of the minimum of  $dU_g$ . Glottis begins to open at time  $t=0$  and begins to close at time  $t_e$ . The time points of  $U_0$  and  $E_e$  are  $t_p$  and  $t_e$  respectively. The time period between  $t_e$  and the projection of the tangent of  $dU_g$  at  $t_e$  is  $t_a$  and  $t_n$  is equal to  $t_e - t_p$ . OQ is the ratio of open time to total period duration, i.e.  $OQ = (t_e - 0) / t_p$ . SQ is the ratio of the duration of the rising portion to the duration of the falling portion of the glottal open phase, i.e.  $SQ = t_p / (t_e - t_p)$ . OQ influences the relative energy level of the first harmonic [2] while SQ is related to the energy level of the first, second and third harmonics [10][14]. Increasing TL (spectral Tilt Level) attenuates the high-frequency components associated with "corner rounding" resulting from the non-simultaneous closure along the length of the vocal folds [2].

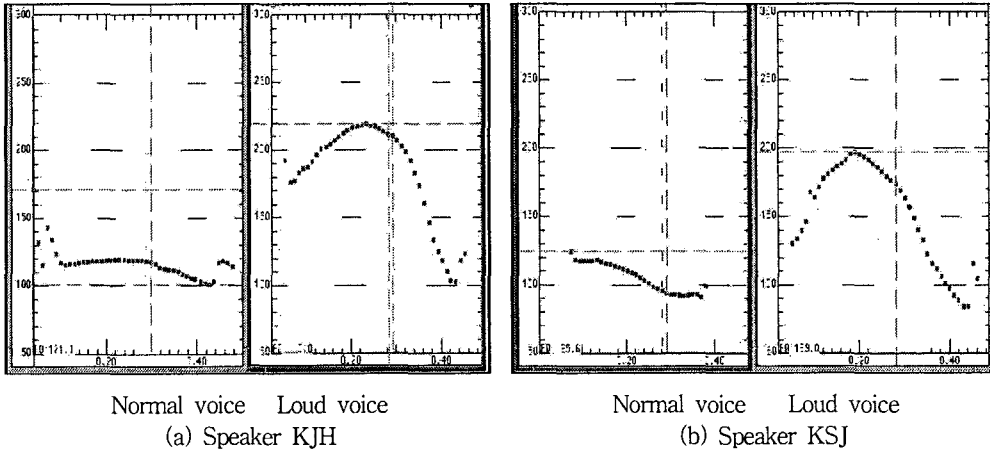


Figure 1. F0 contours of [a] vowel in loud and normal voices (horizontal axis represents time in seconds and vertical axis represents F0 values in Hz, respectively)

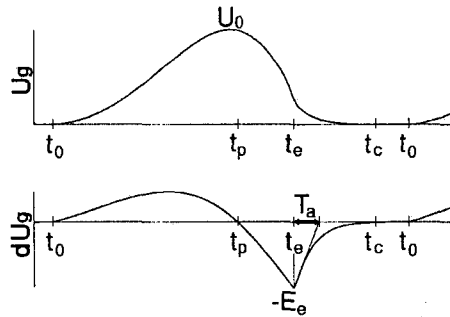


Figure 2. Glottal flow ( $U_g$ ) and glottal flow derivative ( $dU_g$ ) with the parameters of the LF-model

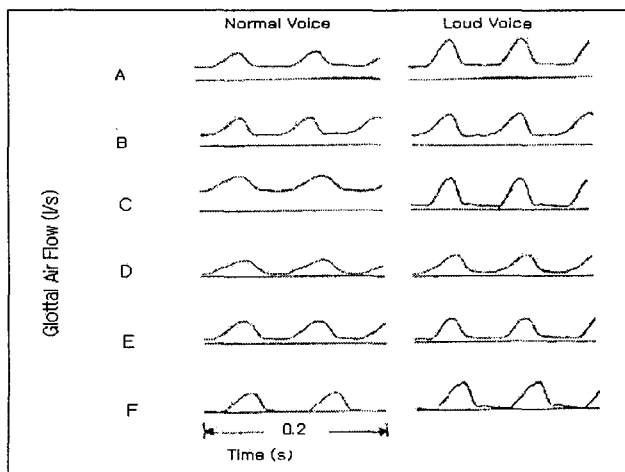


Figure 3. Glottal airflow versus time waveforms in normal, and loud voice for six male speakers [1]

Glottal waveforms in figure 3 are obtained from six male speakers (A, B, C, D, E, F) by Homberg et al. As can be seen in figure 3, the glottal waveforms of loud voices show sharp angles between the end of the closing and the beginning of the closed portions [1], which result in the boost of the high frequency components, suggesting the decrease of TL [2]. The closed portion in a loud voice is often well defined and sometimes relatively longer, which would result in a smaller OQ [1]. SQ in a loud voice is greater than the one in a normal voice [1]. Greater SQ means that the glottis closes more rapidly. According to research where twenty-five male speakers' voices in loud and normal phonations were measured, the average OQ decreased and the average SQ increased from normal to loud phonation and loud voices were typically produced with higher fundamental frequency than with normal voices [1].

Spectrums in figure 4 were obtained from one of ten participants whose spectrums revealed similar patterns. In figure 4 (b), increased bandwidth (the width of harmonics) shows increased F0 for the loud phonation. Spectrum tilt level (TL) in figure 4 (b) decreased. The first harmonic component becomes weaker in a loud condition suggesting smaller OQ [2].

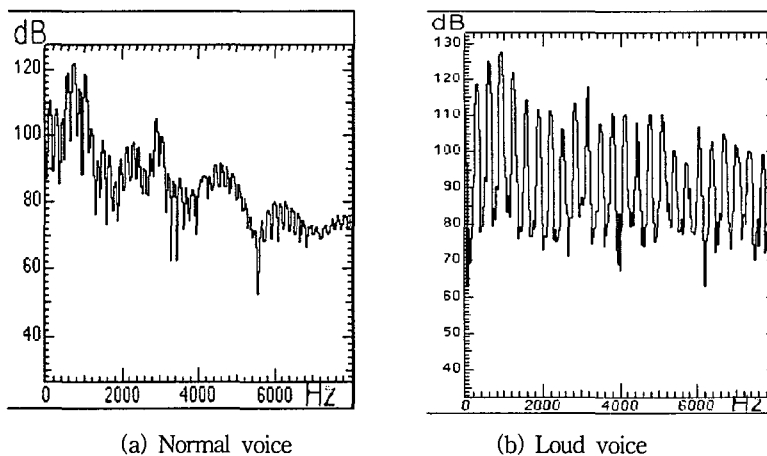


Figure 4. Spectrum of [a] phonation by a male in normal and loud voices

### 3. Experiment

We synthesized the [a] vowels by using the average first, second and third formant values from 76 male speakers [16]. We divided the four parameters (F0, SQ, OQ, TL) into three levels. Table 1 shows F0 values according to each level. Five values for each level indicated each vertex of piecewise linear curves (see figure 5).

F0 contour extracted from natural voices had three levels of forms (see figure 5).

Piecewise linear curves were used to approximate the original F0 curve. Piecewise linear curve and the original F0 curve showed little perceptual differences according to the preliminary listening test. To avoid the argument that piecewise linear curves are types, not values, we computed F0 tilt values by using formula (1). Formula(1) is based on the study that the dynamic component of F0 contours influences the perception of loudness [15].

$$\text{F0 tilt (Hz/ms)} = \text{Max\_F0} - \text{Initial\_F0} / \text{T2} - \text{T1} \quad (1)$$

In formula (1), initial\_F0 is the starting value of the F0 contour and Max\_F0 is the maximum value of the F0 contour, T1 and T2 are the time values of Initial\_F0 and max\_F0 respectively (e.g. F0 tilt for level 1 = (156 Hz - 136.2 Hz) / (285 ms - 25 ms)). Equation (1) is based on the study that the dynamic component of F0 contour influences the perception of loudness [15]. The greater the F0 tilt, the more dynamic the contour becomes.

Table 1. F0 values for each level

	Values of vertexes of piecewise linear F0 curve				
	(25ms)	(155ms)	(285ms)	(500ms)	(600ms)
Level 1	136.2 Hz	146 Hz	156 Hz	146 Hz	102 Hz
Level 2	159.5 Hz	220.3 Hz	237.5 Hz	223 Hz	102 Hz
Level 3	182.8 Hz	294.6 Hz	319 Hz	300 Hz	102 Hz

Table 2. Parameter values at three levels

	F0 tilt	SQ	OQ	TL
Level 1	0.076	1.32	4.6	0 dB
Level 2	0.300	2.19	6.2	6 dB
Level 3	0.524	3.06	7.8	12 dB

OQ values and SQ values at levels one and three are the minimum and maximum values of OQ and SQ of the 25 male speakers' glottal waveforms [1]. OQ values and SQ values at level two are the averages of levels one and three values. TL values are divided into 12 dB, 6 dB and 0 dB.

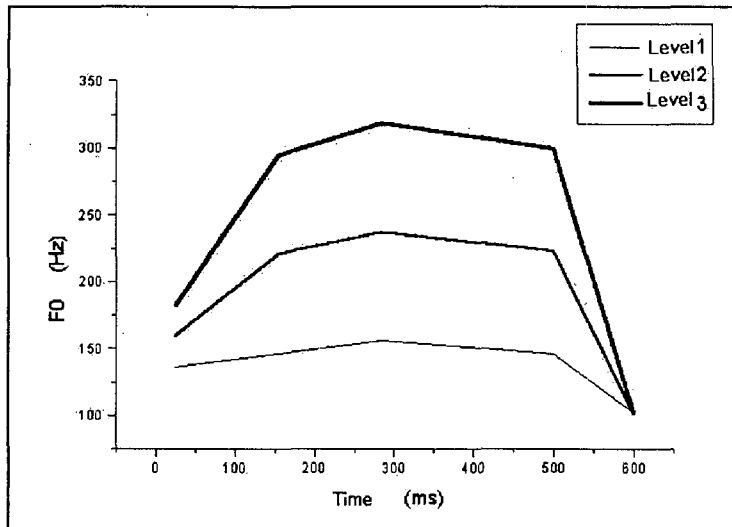


Figure 5. Three F0 curves used to make [a] vowels

F0 tilt, SQ values at level one and OQ and TL values at level three with the rest of the parameters at default values are used to make a normal sound [3]. We increased the AV (Amplitude of Voicing) of the sound from 60 dB to 70dB (maximum AV is 80 dB) to make a reference sound labeled A. In doing this, we intended to show that amplitude variation is not the main effective factor for the perception of loud voice. If amplitude variation is a dominantly contributing factor for loud voice perception, reference sound A which has increased amplitude, will be enough for loud voice perception.

Reference sound B is made up of F0 contour at level 3, SQ at level 3, OQ at level 1 and TL at level 1. Reference sound B has the acoustic characteristics of loud voices in natural sound. According to Humbert et al., loud voices have smaller OQ and TL and larger F0 and SQ values than normal voices [1]. Listeners participating in a preliminary listening test came to the unanimous consensus that reference sound B is louder than reference sound A. The rationale behind reference sounds A and B is to avoid the psychological phenomena of subject tendency to compromise the estimation at certain values. Thus, listeners can anchor their ratings.

Each of the four parameters are divided into 3 levels to make eighty-one ( $3 \times 3 \times 3 \times 3$ ) combinations of synthetic sounds. All of the synthetic sounds are energy normalized to compensate for the energy difference among the sounds. Loud voices in natural sound tend to have a shorter duration [1]. This paper, however, doesn't take this 'shortness of duration' phenomena into consideration.

Figure 6 depicts the computer interface used for the listening test. The computer interface used in this study adopted an analogue scale method. Compared with a discrete scale method, an analogue method enhances the consistency of the listeners' judgement [5].

With this interface, the whole stimuli can be seen and compared easily [5]. Listeners can also find a specific stimulus with ease and listen to it as many as they want. [5].

Eighty-one slide bars can be seen, numbered from 1 to 81. To the right of each bar is a box displaying the value which changes according to the position of the button on the bar. If the button on the left of each bar is clicked, one of [a] sounds assigned to it can be heard. Above the bars are buttons inscribed with Min (0%) and Max (100%). If the button with Min (0%) or Max (100%) is clicked, reference sound A or reference sound B can be heard. Listeners are supposed to compare eighty-one sounds with reference sounds A and B. Then they also compare the eighty-one sounds with each other. Listeners then judge how much louder a sound is than the reference sounds by dragging the button on the bar. If a sound is louder than reference sound A, the button on the bar should be moved to the right of the button assigned for reference sound A. The louder a sound is, the farther the button should be moved to the right. The perceptual values getting out of the range from 0% to 100% are covered in the range from -100% to +200%. The four markers on each bar indicate positions for -100%, 0%, +100% and +200% from left to right.

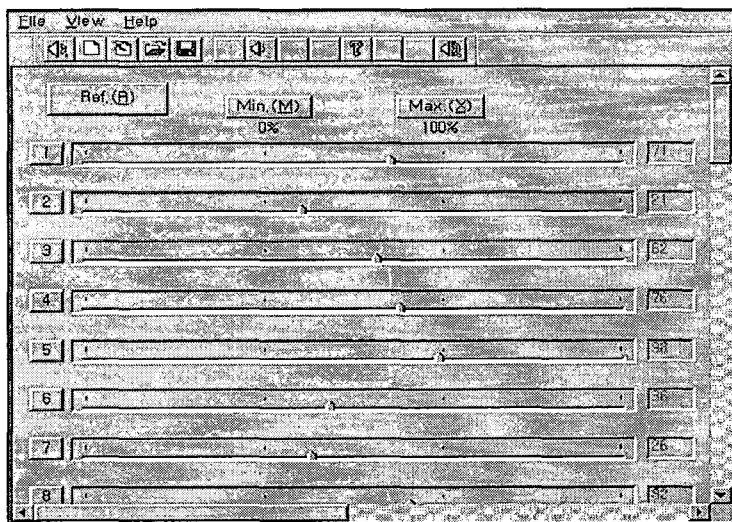


Figure 6. Computer interface used for the listening test

For example, if a button assigned to a sound has the same perceptual loudness as reference sound A, listeners are supposed to drag the button on the bar to the position of 0% where the 'Min' button is located. The position of each button is determined by the relative loudness in comparison with the reference sound. The more similar the degree of perceptual loudness between sounds, the shorter the distance between the buttons becomes.

Two kinds of test are conducted. The first test is done with four parameters (F0\_tilt, SQ, OQ and TL) of three levels changing to make eighty-one combinations. Ten listeners

without any hearing problem participate in the tests. All of them are in their twenties and not familiar with the synthesized sounds being evaluated. In this way, we hope to avoid any influence on the results based on previous experience. Sounds are sorted at random each time. Two sounds which have the same loudness as reference sounds A and B are also included in the eighty-one sounds to be tested. Some of the data are excluded when a listener judges these as different sounds from the reference sounds, since the validity of the judgement is questionable.

#### 4. Analysis of the Experiments

##### 4.1 First Experiment

The four parameters, F0 tilt, SQ, OQ and TL are used to make eighty-one combinations ( $3 \times 3 \times 3 \times 3$ ) in the first experiment. F values and P values are obtained from the multi-way factorial design of the first experiment (see table 3). The P values of the interactions among the parameters are less than 1%, which is statistically insignificant. F0 tilt, SQ, OQ and TL are statistically significant.

Table 3. F and P values obtained from factorial design analysis of eighty-one sounds

	F value	P value
TL	5.85	<u>0.0031</u>
OQ	10.69	<u>0.0001</u>
SQ	64.10	<u>0.0001</u>
F0_TILT	1318.78	<u>0.0001</u>
TL*OQ	0.40	0.8073
TL*SQ	0.46	0.7622
TL*F0_TILT	0.15	0.9632
OQ*SQ	1.13	0.3399
OQ*F0_TILT	1.24	0.2934
SQ*F0_TILT	1.42	0.2258

Table 4. Partial correlation coefficients and P values obtained from eighty-one sounds with the perception of loud voice

	Partial correlation coefficients	P values
TL	0.08642	<u>0.0141</u>
OQ	-0.15205	<u>0.0001</u>
SQ	0.33244	<u>0.0001</u>
F0_Tilt	0.87088	<u>0.0001</u>



We used partial correlation coefficients to judge the trend and linearity between the perception of loud voices and other parameters (see table 4). According to table 4, the relationship between the perception of loud voices and other parameters are statistically significant at level 10%. F0 tilt has the highest linearity with the perception of loud voices. SQ has the second highest linearity with the perception of loud voices, OQ the third and TL the fourth.

According to the partial correlation, F0 tilt, SQ and TL have a positive relationship with the perception of loud voice and OQ has a negative relationship. (The greater the value of F0 tilt, the greater the perception of loud voice becomes.)

Multiple regression analysis is computed to evaluate the relative contribution of each parameter to the perception of loud voice. Multiple regression analysis reveals that all parameters are statistically significant (see table 5). Contribution to the perception of loud voice by the parameters are estimated (see table 5). According to the amount of contribution, F0 tilt (73%), SQ (2%), OQ (0.5%) and TL (0.1%) can be listed in descending order. It can be said that F0 tilt is the dominating contributor to the perception of loud voice. Therefore listeners judged the loudness by using the acoustic cue of F0 tilt.

Table 5. Multiple regression analysis of the eighty-one sounds

	Coefficients of multiple regression	Coefficients of determination	F values	P values
TL	0.29	0.0018	6.06	<u>0.0141</u>
OQ	-1.91	0.0055	18.93	<u>0.0001</u>
SQ	8.05	0.0289	97.24	<u>0.0001</u>
F0_Tilt	96.88	0.7310	2195.49	<u>0.0001</u>

Following is the formula of the multiple regression analysis. Standard deviation is in the parentheses.

$$LD = -0.54 + 0.28 TL - 1.91 OQ + 8.05 SQ + 96.88 F0\_Tilt \quad (2)$$

(0.49) (0.12) (0.44) (0.81) (1.93)

It can be said that three parameters (SQ, OQ, TL) are deemphasized because of the dominant influence of F0 tilt on the perception of loud voice. Therefore, there is a need for another experiment with the three parameters alone. To do another experiment, we had to fix the level of F0 tilt at each level. Each level of F0 tilt is combined with twenty-seven combinations of the three parameters (3x3x3), which means three groups of twenty-seven combinations. We had to avoid doing the same kind of experiments because the three groups of twenty-seven combinations are basically the same as the first experiment, except for the F0 tilt level condition. We also had to avoid listening fatigue from too many

listening tasks. According to the multiple regression analysis of the three groups, the group with the F0 tilt at level two has the smallest partial R square, which is a measure for accountability (see table 6). Therefore we can say that any statistically significant outcome from this group guarantees the statistically significant outcomes from the other groups.

Table 6. Coefficients of determination and P values with F0 tilt fixed at each level (in (a) and (b) TL is not statistically significant at 10%.)

(a) F0 tilt fixed at level 1

	Partial R Square	P value
OQ	0.0231	<u>0.0001</u>
SQ	0.1238	<u>0.0076</u>

(b) F0 tilt fixed at level 2

	Partial R Square	P value
OQ	0.0187	<u>0.0201</u>
SQ	0.0667	<u>0.0001</u>

(c) F0 tilt fixed at level 3

	Partial R Square	P value
TL	0.0106	0.0651
OQ	0.0204	<u>0.0112</u>
SQ	0.1447	<u>0.0001</u>

To find the optimal combination out of eighty-one combinations contributing to the perception of loud voices, the sum of values are computed at each level (see table 7). The optimal combination is composed of each level of maximum values since there are no interactions between parameters. According to table 7, therefore, TL at level 2 (6dB), OQ at level 1 (0.46), SQ at level 3 (3.06) and F0 tilt at level 3 (third F0 tilt) make the optimal combination contributing to the perception of loud voices. This agrees with the fact that OQ has a negative correlation with the perception of loud voices and F0 tilt and SQ have positive correlations.

Another way to find the optimal combination is to simply find the combination with the maximum test value out of the eighty-one combinations. The combination of TL at level 3 (6dB), OQ at level 1 (4.6), SQ at level 3 (3.06) and F0 tilt at level 3 (0.851) turned out to be the optimal combination. However, the combination found by this method can be influenced by extreme values, therefore; we came to the conclusion that the first method is more reasonable than the second one to find an optimal combination.

Table 7. Sum of values at each level

	TL	OQ	SQ	F0 tilt
Level 1	13923	<u>15597</u>	12219	5457
Level 2	<u>15135</u>	14369	15691	13981
Level 3	14854	13946	<u>16002</u>	<u>24474</u>

#### 4.2 Second Experiment

We attempted to determine the perceptual effect of the parameters without the influence of F0 tilt since F0 tilt's dominant influence is believed to mitigate the significant difference between the three parameters: SQ, OQ and TL. We made twenty-seven combinations of sounds (3x3x3) out of the three parameters. F0 tilt was fixed at level 2, which means F0 value did not vary.

The interaction between parameters is not statistically significant at 1% (see table 8). SQ (259.23), OQ (22.25), TL(21.09) are statistically significant at 1%.

Table 8. F and P values obtained from factorial design analysis of twenty-seven sounds

	F Value	P Value
TL	21.09	<u>0.0001</u>
OQ	22.25	<u>0.0001</u>
SQ	259.23	<u>0.0001</u>
TL*OQ	0.79	0.5351
TL*SQ	2.46	0.0459
OQ*SQ	2.21	0.0679

The partial correlation between SQ, OQ, TL and the perception of loud voice is statistically significant at 1% (see table 9). Unlike the first experiment, the partial correlation between SQ and OQ, TL, which is believed to be influenced by the exclusion of the dominant factor, F0, is statistically significant at 1%. According to the factorial design analysis of twenty-seven sounds, SQ and TL have a positive relationship with the perception of loud voice, whereas OQ has a negative relationship as it did in the first experiment.

According to the multiple regression of twenty-seven sounds, SQ, OQ and TL are statistically significant at 1%. The coefficients of determination say that SQ has the most contribution: SQ (52.81%), OQ (3.46%) and TL (3.47%). It is believed that the listeners judged the loudness of twenty-seven sounds by using the acoustic cue of SQ mainly without any acoustic cue of F0.

Table 9. Partial correlation coefficients and P values obtained from the twenty-seven sounds with the perception of loud voice

	Partial correlation coefficients	P values
TL	0.28181	<u>0.0001</u>
OQ	-0.28143	<u>0.0001</u>
SQ	0.75329	<u>0.0001</u>

	Partial correlation coefficients	P values
TL	-0.21229	<u>0.0005</u>
OQ	0.21200	<u>0.0005</u>

	Partial correlation coefficients	P values
TL	0.07931	0.1955

Table 10. Multiple regression analysis of the twenty-seven sounds

	Coefficients of multiple regression	Coefficients of determination	F values	P values
TL	1.89	0.0347	21.21	<u>0.0001</u>
OQ	-7.08	0.0346	22.88	<u>0.0001</u>
SQ	50.82	<u>0.5281</u>	299.91	<u>0.0001</u>

To find the optimal combination out of twenty-seven sounds contributing to the perception of loud voices, the sum of values are computed at each level (see table 11). Again, the optimal combination is composed of each level of maximum value since there are no interactions between parameters. According to table 11, therefore, TL at level 2 (6dB), OQ at level 1 (0.46) and SQ at level 3 (3.06) make the optimal combination contributing to the perception of loud voices the same as the first experiment. This agrees with the result that OQ has a negative correlation and SQ has a positive correlation with the perception of loud voices.

Another way to find the optimal combination is to simply find the combination with the maximum test value out of twenty-seven combinations. The combination of TL at level 3 (6dB), OQ at level 1 (4.6) and SQ at level 3 (3.06) turned out to have the maximum test value. However, the combination found by this method can be influenced by extreme values. Therefore, as in the case of the first experiment, the first method is more appropriate than the second one.

Table 11. Sum of values at each level

	TL	OQ	SQ
Level 1	3540	<u>6345</u>	156
Level 2	<u>5655</u>	4124	6505
Level 3	5581	4307	<u>8115</u>

## 5. Discussion

Comparing the sum of values at each level, we found that F0 tilt at level 3, SQ at level 3 (3.06), OQ at level 1 (0.46) and TL at level 2 (6dB) make the optimal combination contributing to the perception of loud voices. The dynamic characteristic of F0 tilt showed a dominant influence on the perception of loud voices. Neuhoff et al. reported that the dynamic characteristic of the F0 contour contributes to the perception of loudness [15]. The influence of SQ is dominant in the second experiment in which the influence of F0 tilt is excluded.

The increase of SQ is known to reduce the energy level of first, second and third harmonics leading to the vocal quality of pressed voices [10][14]. Therefore, the increase of SQ contributed to the perception of loud voices by increasing the auditory effect of pressed voices. While F0 tilt has a direct influence on loudness, SQ, OQ and TL seem to influence the vocal quality found in loud voices.

Rather than comparing various vowels, only /a/ vowels are used in this paper because the categorical influence of vowels has no significant impact on the loudness. The glottal parameters of loudness are mainly influenced by the glottal status, not by the kind of vowels. Glottal condition of a phonation, however, changes at both word and sentence levels. Our next study should expand the range from a single vowel to words and sentences.

## References

- [1] Eva B. Holmberg, Robert E. Hillman and Joseph S. Perkell. 1988. "Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice," JASA, vol. 84, no. 2, pp. 511-529.
- [2] Dennis H. Klatt and Laura C. Klatt. 1990. "Analysis, synthesis, and perception of voice quality variations among female and male talkers," JASA, vol. 87, no. 2, pp. 820-857.
- [3] Dennis Klatt. 1990. "Description of the cascade/parallel formant synthesizer," KLATTALK, Conversion of English Text to Speech, Chapter 3.
- [4] Randall B. Monsen and A. Maynard Engebetson. 1977. "Study of variations in the male

- and female glottal wave," *JASA*, vol. 62, no. 4, pp. 981-993.
- [5] Svante Granqvist. 1996. "Enhancements to the Visual Analogue Scale, VAS, for listening tests," *TMH-QPSR* 4/1996.
- [6] Gunnar Fant, Johan Liljencrants and Qi-guang Lin. 1985. "A four-parameter model of glottal flow," *STL-QPSR* 4/1985.
- [7] Rolf Carlson, Gunnar Fant, Christer Gobl, Bjorn Granstrom, Inger Karsson and Qi-Guang Lin. 1989. "Voice source rules for text-to-speech synthesis," *Proc. IEEE ICASSP* 89, pp. 223-226.
- [8] D. G. Childers and C. K. Lee. 1991. "Vocal quality factors: Analysis, synthesis, and perception," *JASA* vol. 90, no. 5, pp. 2394-2410.
- [9] Park, S. H. 1992. *Modern Factorial Design Analysis*, Minyoungsa.
- [10] H. Strik. 1994. *Physiological control and behaviour of the voice source in the production of prosody*, Katholieke Universiteit Nijmegen.
- [11] H. Strik and L. Boves. 1992 "Control of fundamental frequency, intensity and voice quality in speech," *Journal of Phonetics*, vol 20, pp. 15-25.
- [12] H. Strik and L. Boves. 1993. "A physiological model of intonation," *Proceedings of the Department of Language and Speech, Nijmegen University*, 16/17, pp. 96-105.
- [13] H. Strik and L. Boves, "Downtrend in F0 and P<sub>sb</sub>," *Journal of Phonetics*, (to be published).
- [14] Inger Karlsson. 1990 "Voice source dynamics for female speakers," *Proceedings of the 1990 International Conference on Spoken Language Processing, Kobe*, pp. 69-72.
- [15] J. Neuhoff and M. McBeath. 1997 "The Interaction of pitch and loudness in dynamic stimuli: beyond the doppler illusion," *133rd ASA Meeting*.
- [16] G. E., Peterson and H. L. Barney. 1952. "Control methods used in a study of the vowels used in a study of the vowels," *JASA*, vol. 24, pp. 175-184.
- [17] Stellan Hertegaard. 1994. *Vocal fold vibrations as studied with flow inverse filtering*, Department of Logopedics and Phoniatrics No. 5, Huddinge University Hospital.
- [18] Stellan Hertegaard, Personal communication.

Received: Jan. 30, 2001.

Accepted: Mar 3, 2001.

▲ Yi, Sopae

Department of Cognitive Science

Pusan National University

San 30 Jang-Jun Dong, Gum-Jung Goo, Pusan, 609-735, Korea

E-mail: spyi@web.pusan.ac.kr

▲ Lee, One Good

Department of English Literature and Language

Pusan National University

San 30 Jang-Jun Dong, Gum-Jung Goo, Pusan, 609-735, Korea

E-mail: oglee@web.pusan.ac.kr

- ▲ Kim, Hyung Soon  
Department of Electronics Engineering  
Pusan National University  
San 30 Jang-Jun Dong, Gum-Jung Goo, Pusan, 609-735, Korea  
E-mail: kimhs@hyowon.cc.pusan.ac.kr