Ann. Bull. RILP (1988) No. 22, 75-83

THE EFFECT OF LARYNGEAL SURGERY ON THE SPEECH CEPSTRUM

Yasuo Koike and Junji Kouda

Introduction

For some time it has been known that the cepstral peaks in the acoustic speech waveform are sensitive to the existence of laryngeal pathology[1]. This seems to be because the periodicity of the acoustic speech waveform, reflecting the regularity of the vocal fold vibration, can be affected by various pathologic changes in the larynx. Conversely, if such pathologic changes could be effectively removed from the larynx, then the periodicity of the acoustic signal should be improved, and thus the cepstral peaks should be restored. The removal of benign lesions from the affected vocal fold constitutes an occasion for investigating whether or not such a restoration of the cepstral peaks in fact takes place.

The present study has been undertaken to confirm that such a change attributable to surgery actually occurs in the cepstrum, and to determine the characteristic features of such a change in the cepstral display. Since the purpose of most surgeries for benign lesions in the larynx is to restore a normal voice, an assessment of the change in the cepstrum caused by a surgical operation may in a sense be regarded as a method for evaluating the success of the surgery. Cepstrum analysis may be particularly useful for this type of evaluation since it can be carried out objectively without the intervention of human judgement. The need for objective means for evaluating the effect of surgeries does not require lengthy explanation.

The effect on the speech cepstrum of surgery for laryngeal malignancy was not considered in the present paper, since the objective of such operations may be entirely differnt from the restoration of a normal voice. Even an almost complete destruction of vocal function may result from a successful surgery. The assessment of vocal function with the speech cepstrum may not be relevant to the evaluation of the success of such operations.

DEPARTMENT OF OTOLARYNGOLOGY, THE UNIVERSITY OF TOKUSHIMA SCHOOL OF MEDICINE

Method

The subjects of the present study consisted of 22 patients with various laryngeal diseases. The sexes, ages, and the pathologies based on histological examinations are listed in Table 1. The subjects were instructed to sustain the vowel /a/ for approximately 4 seconds at a comfortable pitch and loudness. A magnetic tape recording of the acoustic speech was made for each subject with an electret condenser microphone (Sony ECM-969) and a cassette-tape recorder (Sony TC-D5M). Recordings were made before and after the treatment for each patient. Although surgery was performed on neither case #21 (glottic cancer) nor case #22 (glottic tuberculosis), the data from these patients were also included as a refernce.

A stable portion of the vowel of about 1 second in duration was adopted for analysis. These materials were digitized with an analog to digital converter at a sampling rate of 10 KHz with the aid of a computer (PDP-11/73). An anti-aliasing filter was applied prior to the A/D conversion. The digitized data were stored on a floppy disk for each subject. The data were then analyzed with a cepstrum implemented on the computer. A Hamming window having a duration of 40 msec was employed. The cepstrum is the power spectrum of the log spectrum of the speech wave, and is widely adopted for pitch detection. A detailed description of this analysis procedure is available elsewhere[2),3)]. Five cepstra from five non-overlapping intervals were made for each utterance, and the one showing the most explicit periodicity was selected and employed for subsequent study.

Twenty cases out of the whole subject group underwent laryngo-microsurgery under general anesthesia. The method of surgery employed varied from patient to patient, depending upon the type of pathology in the larynx. Polyps or nodules were removed with conventional procedures. A "squeezing technique"[4)] was adopted for patients with polypoid degeneration. A laser surgery was performed on the patient with papilloma. As was mentioned earlier, the patient with glottic cancer was given radio-therapy, and no surgery was performed. The patient with glottic tuberculosis was treated with antituberculous agents, and no surgery was attempted.

Results

The most remarkable effect of laryngeal surgery on the cepstra for these patients seemed to be a lengthening of the

Table 1

The sex, age, and the type of laryngeal pathology for each subject.

	case	sex	age	pathology
1	C. I.	f	19	vocal nodule
2	S. T.	m	59	*id.*
3	M. A.	m	50	vocal polyp
4	K. K.	m	38	*id.*
5	T. M.	f	44	*id.*
6	N. S.	m	42	*id.*
7	T. T.	m	73	*id.*
8	S. N.	m	57	*id.*
9	T. B.	f	41	*id.*
10	R. S.	m	61	*id.*
11	N. I.	m	63	polypoid degeneration
12	M. T.	m	62	*id.*
13	K. I.	f	56	*id.*
14	A. K.	f	48	*id.*
15	C. M.	f	53	prolapsus ventriculi
16	M. T.	m	42	ventricular polyp
17	M. O.	m	51	laryngeal papilloma
18	Y. K.	m	39	laryngeal granuloma
19	K. S.	m	60	*id.*
20	F. S.	m	58	*id.*
21	S. M.	m	78	glottic carcinoma
22	T. S.	m	48	glottic tuberculosis

dominant cepstral peak. This effect is illustrated in Figure 1, which is based on the data obtained from a patient with vocal polyp (case #10). The dominant peak observable in the cepstrum after the surgery (right tracing) is seen to be much taller than that for the speech wave recorded before the surgery (left tracing). The abscissa of this Figure shows "quefrency" in msec., and the ordinate is cepstral energy plotted on an arbitrary scale.

It is interesting to note that this type of change in the cepstral peak corresponded well with the improvement in the perceived hoarseness of the patient's voice. In other words, there was invariably an explicit lengthening of the dominant cepstral peak, for the cases showing an apparent improvement in the perceived voice quality after the surgery. This type of cepstral change was not quite remarkable, however, for patients who had had a very low degree of hoarseness to begin with. The cepstral peaks for these patients had been generally quite tall even before the surgical intervention, and there was little room for improvement.

A similar elongation of the dominant cepstral peak was observed after surgery in the rest of the patients with vocal polyp or vocal nodule, though the degree of lengthening varied from patient to patient. The subjects with other benign lesions, such as papilloma or granuloma, also revealed an elevation of the dominant cepstral peak after the surgical treatments, including laser surgery. Here again, the extent of the cepstral change varied greatly from patient to patient, depending mainly upon how much the vibratory behavior had been affected by these lesions before the surgical operation.

In addition to the lengthening of the dominant cepstral peak, an apparent shift of the fundamental frequency was generally observed in patients with polypoid degeneration of the vocal fold. Figure 2 shows this effect on the cepstral display. The dominant cepstral peak in the right tracing, which denotes the periodicity of the acoustic waveform after the surgery, is not only taller than that in the left tracing, but also considerably shifted to the left side, indicating an increase in the fundamental frequency. For this subject, the fundamental frequency was 74 Hz before surgery and 104 Hz after.

Figure 3 shows the corresponding results for case #21 (glottic carcinoma) which involved radiation therapy instead of surgery. Here also we see a remarkable improvement in the regularity of laryngeal vibration. The dominant cepstral peak after the treatment (right tracing) is apparently taller than

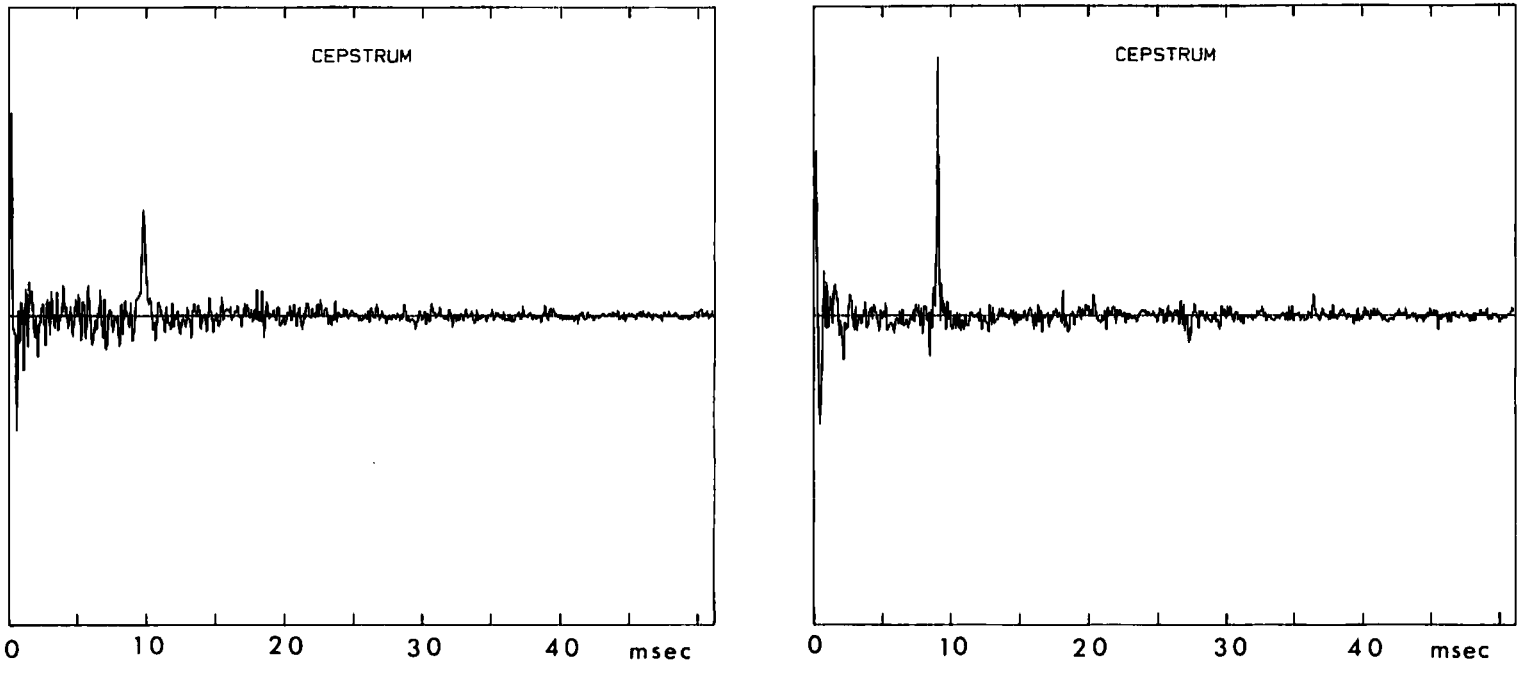

Figure 1. The cepstra of the acoustic speech of a patient with vocal polyp (Case No. 10). The left tracing reveals the cepstral display of the signal obtained before the surgery, and the right tracing the same after the surgical operation. An apparent lengthening of the cepstral peak can be observed after the treatment. The abscissa shows "quefrency" in msec., and the ordinate indicates an arbitrary scale.

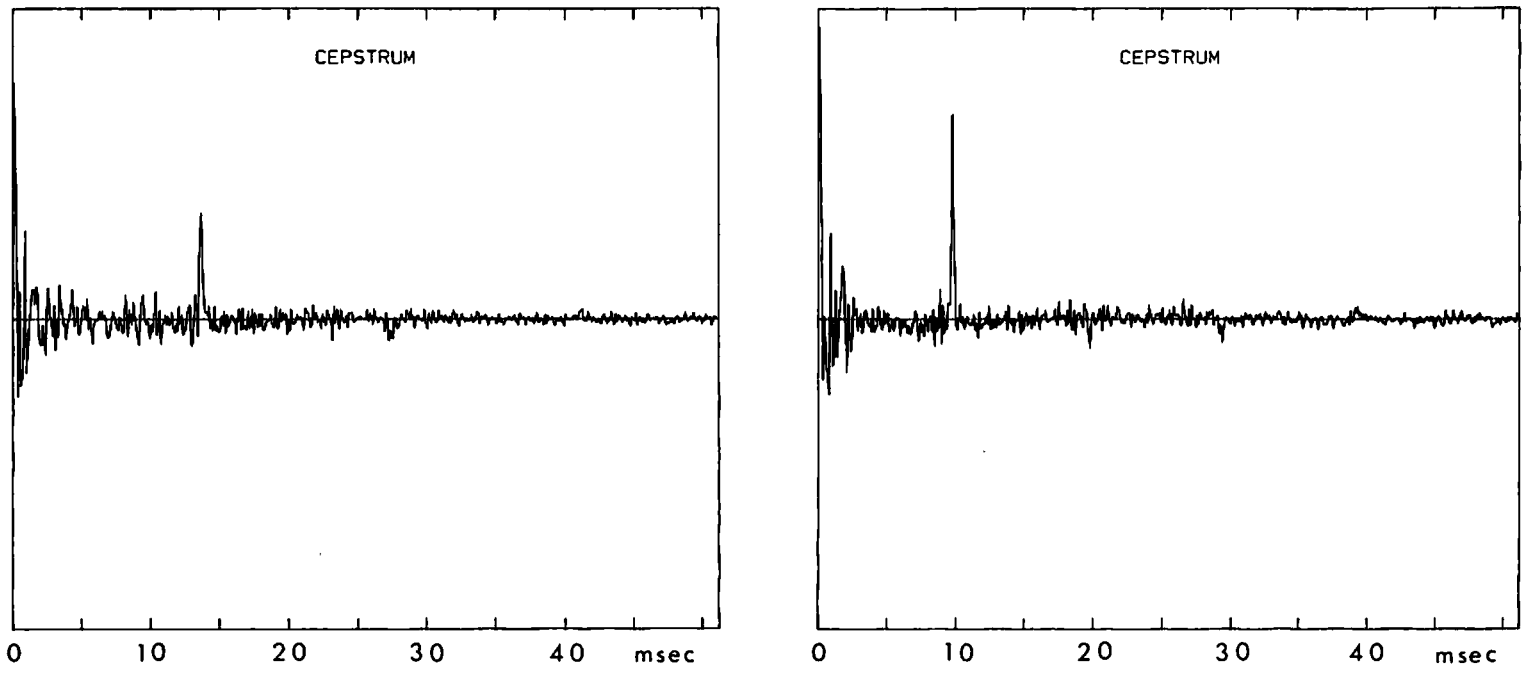

Figure 2. The cepstra of the acoustic speech of a patient with polypoid degeneration of the both folds (Case No. 11). The left and the right tracings reveal the cepstrum before (left tracing) and after (right tracing) the operative intervention respectively. The dominant cepstral peak is not only lengthened, but also obviously shifted to the left side on the quefrency scale after the operation.

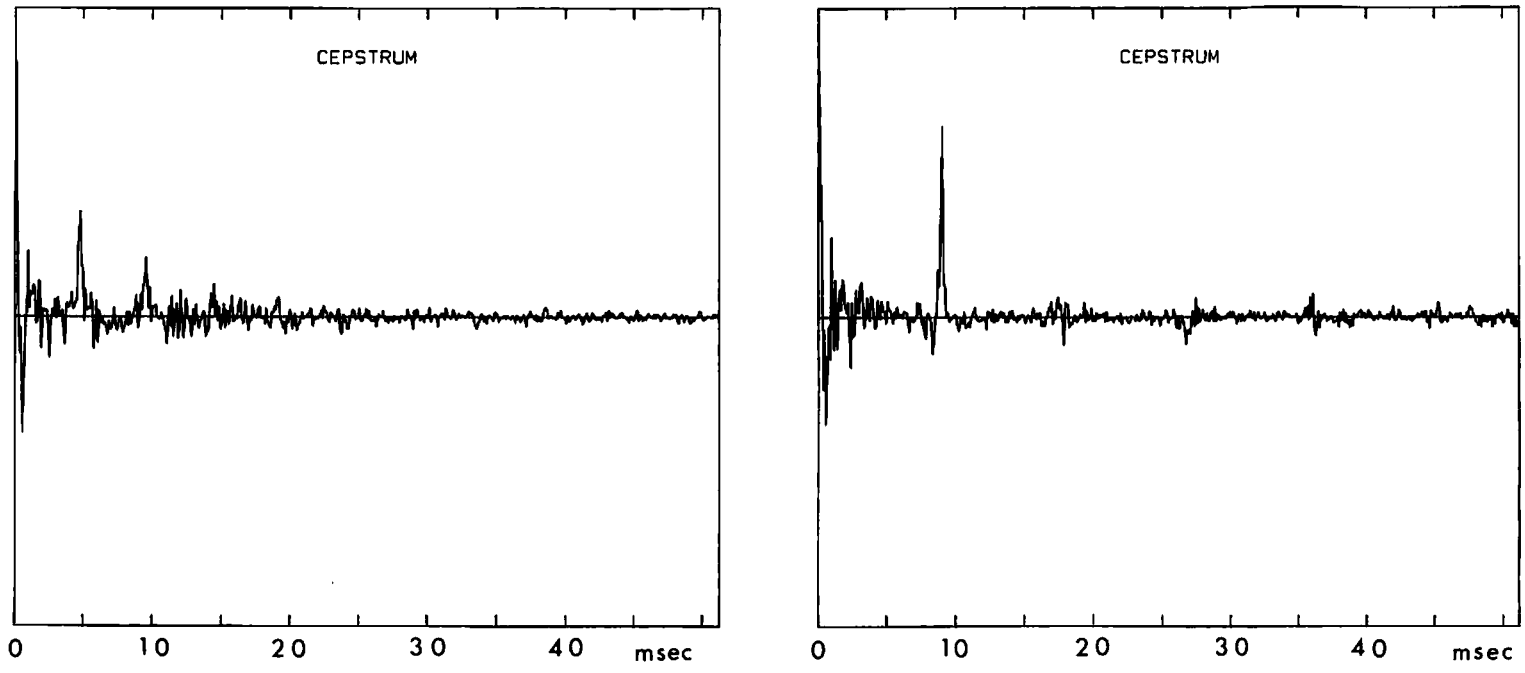

Figure 3. The cepstra of the acoustic speech of a patient with glottic cancer (Case No. 21). The left tracing demonstrates the cepstral display of the acoustic signal before the radiation therapy, and the right tracing discloses the same after the treatment. The dominant peak in the right tracing is shifted to the right side on the quefrency scale, indicating a downward shift of the fundamental frequency after the therapy.

that before the therapy (left tracing). The tumor which had existed on the vocal fold completely disappeared after the irradiation, though some edema of the fold remained. The perceived voice quality of this patient was improved, and the cepstral change seemed to represent the restoration of regular vibratory motion of the fold fairly well. In this particular case, in addition, a lowering of the fundamental frequency from 216 Hz to 111 Hz was also noted following the treatment.

Case #22 with glottic tuberculosis also revealed an elongation of the dominant cepstral peak after the administration of antituberculous agents. With this case also, the change in the cepstral peak corresponded well with an improvement in the voice quality.

Discussion

Our data presented above display an appreciable difference in the dominant cepstral peak, attributable to laryngeal surgery. In most cases with benign lesions on or surrounding the vocal fold, an apparent lengthening of the dominant cepstral peak is disclosed following the surgery. This indicates that the periodicity of the acoustic speech waveform has been improved after the surgical operation, implying that the regularity of the vocal fold vibration has been restored. Since this type of change in the cepstral display is particularly clear when the pre-operative vibratory behavior in the larynx has been remarkably distorted, the cepstrum analysis seems to be especially suited for the evaluation of the effect of surgery in moderate or severe cases with benign lesions.

The fact that the elongation of the dominant cepstral peak generally corresponds to an improvement in the perceived voice quality also seems to be noteworthy. Although the regularity of the speech signal can be revealed by the height of the dominant cepstral peak, the scale on the ordinate of the cepstral display is not linear. This makes quantitative analysis of the length of the peaks rather complicated. Quantitative measurement of the height of the cepstral peaks has not been attempted in the present study for this reason. The investigation of the relationship between the elevation of the dominant cepstral peak and the degree of perceived hoarseness, nevertheless, seems to be a worthwhile topic to pursue, since the cepstrum analysis can at least be performed objectively without human judgement.

The apparent upward shift of the fundamental frequency,

found in the cases with polypoid degeneration, also seems to be an interesting finding. Although clinically it has been known that these patients in general have a lowered fundamental frequency, the post-operative recovery has not yet been sufficiently studied, mainly because of the difficulty in determining the fundamental frequency. Because the cepstrum analysis was originally designed to measure the fundamental frequency when it exists, it is an easy task to observe this shift on the cepstral display. This feature of the cepstrum seems to be convenient for measuring a downward shift of the fundamental frequency as well, as demonstrated in Figure 3. This type of downward shift, however, seems to be peculiar to this case, since the fundamental frequency for patients with laryngeal cancer is in general lowered, and it is rather rare to find an elevated fundamental frequency as was the case with this patient.

Summary

Cepstrum analysis of the acoustic speech waveform was performed on 22 patients with various benign laryngeal lesions, both before and after treatment. In most cases with laryngeal polyp or nodule, a lengthening of the dominant cepstral peak was observed following surgery. This elongation was in agreement with the improvement in the perceived hoarseness. An upward shift of the fundamental frequency was revealed in patients with polypoid degeneration, in addition to the elevation of the dominant cepstral peak above mentioned. It was considered that the cepstrum analysis may be useful in evaluating the success of surgeries for benign lesions on or around the vocal fold.

References

1) Koike, Y. : Cepstrum analysis of pathologic voices. J. of Phonetics 14, 501-507, 1986.
2) Noll, A. M. : Short-Time Spectrum and "Cepstrum" Techniques for Vocal-Pitch Detection. J. Acoust. Soc. Am. 36, 296-302, 1964.
3) Oppenheim, A. V. : Speech Analysis-Synthesis System Based on Homomorphic Filtering. J. Acoust. Soc. Am. 45, 458-465, 1968.
4) Koike, Y. : Laryngomicrosurgery. in Illustrated Handbook of Otolaryngological and Head and Neck Surgery Vol. 2, 88-93, Medical View, Tokyo, 1986.