PHONETIC ASPECTS OF LARYNGECTOMEE SPEECH

Lennart Nord
Department of Speech Communication and Music Acoustics,
Royal Institute of Technology (KTH), Box 70014,
S-100 44 Stockholm, Sweden
and

Britta Hammarberg
Department of Logopedics and Phoniatrics,
Huddinge Univ Hospital,
S-141 86 Huddinge, Sweden

INTRODUCTION

The alternative speech methods that are available today for the laryngectomized have each one its advantages and disadvantages. and each method seems to have its proponents. It is important to investigate the features of the different techniques and also to get a deeper insight into what signifies an acceptable voice following laryngectomy. Moreover, the basic mechanisms of the different kinds of alaryngeal phonation are not very well known, and a comparison with the function of the normal glottis is therefore of interest.
We are presently investigating three different techniques: (i) esophageal speech, (ii) tracheo-esophageal speech (TE-speech) where part of the esophagus is used as voice source, and the driving pressure is generated by the lungs via a valve and finally (iii) electrolaryngeal speech, using an electromechanical device (Servox) that emits pulses through the pharynx wall. For a survey of the area, see [1,2].

STUDIES OF ALARYNGEAL SPEECH

The introduction of the tracheo-esophageal technique has in many ways implied a significant improvement for the establishment of a new voice after laryngectomy. A number of recent studies have compared the different types of alaryngeal speech. Speech intensity, intonation range, intelligibility and listener acceptability have been investigated [3-5], as well as prosodic aspects of the different speaking methods [6,7]. We will present some preliminary findings of acoustic-phonetic analysis of esophageal and tracheo-esophageal speech. Parts of this project have been presented earlier [8,9]. For more extensive reports, that will also cover some clinical aspects, see [10,11].

SPEECH MATERIAL

The speech material was designed with the aim of testing prosodic aspects of communication. Sentences with contrastive word emphasis and short questions and statements typical of a dialogue were thus included. A few vowels were embedded in carrier phrases and the subjects were asked to produce them with low, medium and high intensity. A reading of a short paragraph gave us the material to extract a long term average spectrum of the voice as well as establishing the pitch range of the subject. To investigate air pressure and air flow conditions and also to get an estimate of the voice source characteristics, a flow mask [12] was used in a separate reading of nonsense words.

SPEAKERS

Presently, we have recorded three male esophageal speakers and three male TE-

speakers. A few laryngeal speakers will also be recorded and used as reference. So far, only one male speaker of the same age group has been analysed. Our intention is to collect a representative amount of speakers, that are using the different techniques and also record some multi-technique speakers.

ANALYSIS

The acoustic analysis includes measurements of pitch, intensity, duration, speaking rate, long time average spectral shape and voice source characteristics, derived by inverse filtering. Detailed spectral analysis of articulation and source features are also performed at selected points.
As an important part of the study the acoustic measurements will be correlated to perceptual evaluations, using different groups of listeners.

Overall speech quality

The speech quality differs a lot between the speakers in our material. In Figure 1 spectrographic illustrations of two of the speakers are shown. As can be seen the esophageal voice looks quite normal (top), the deviant quality is here mainly perceived in the prosodic pattern and the low pitch. The TE-speaker has a very harsh voice (bottom) with irregular formant excitations and superimposed noise.

Pitch and timbre

Pitch extraction. Pitch extraction turned out to be difficult and uncertain for these deviant voices. The reasons for this seem to be that the pitch values are usually extremely low and sometimes very aperiodic. Moreover, in contrast to the case of a laryngeal low pitched voice, inverse filtering reveals that the instant of excitation in a tracheo-esophageal voice is not well defined. In Figure 2 a short interval of the vowel /a/ is shown together with the extracted voice source pulse, for a tracheoesophageal speaker, who is using a Panje-valve (note the difference in time scale).
In the frequency domain the fundamental (F0) is very weak relative to the formants, which means that a pitch extraction algorithm based on a spectral measure will work poorly.

Voice source spectrum

When changing from strong to weak voice a typical feature of laryngeal voices is that the fundamental will dominate the spectrum. This is not the case for these voices. Instead, the fundamental will maintain a low level irrespective of the total intensity. This lack of dynamics probably contributes to the unnatural quality of the voice. The perceptual dimension *hyperfunctional*, or tense voice also correlates well with the spectral dimension of relatively weak fundamental [13].

INTONATION

In the production of voiced sounds there are no intricate laryngeal-like mechanisms to stretch and manipulate the vibrating parts in the esophagus. Although a pseudo-glottis is developed, we know little about its function, apart from the fact, that for some speakers it will serve as a surprisingly good substitute.
The analysis of the sentences containing emphatic word stress was made manually, using a speech wave editing program, developed at the Speech Communication Department. Each subjectively judged periodic cycle was marked and the pitch was automatically calculated. Figure 3 shows a set of emphatic sentences *Gunnar målade dörrarna gröna* for one of the TE-speakers, where the word stress was moved one step to the right for each display a) to d). Two things are worth noticing. First, the mean F0 value is rather low for this speaker. Second, the signalling of emphasis seems to be

reasonably well managed by the speaker, in these plots an F0 hump is seen moving to the right for each new sentence (the approximate location of the emphasized word is underlined).

REFERENCES

[1] Shanks,J (Editor) (1986) "Current Strategies of Rehabilitation of the Laryngectomized Patient" Seminars in Speech and Language, Vol 7, No 1, February 1986

[2] Hammarberg, B.(1987) "Erfarenheter av rehabilitering med Panjeventil." Lary nr 2/3 1987, sid 16-20 (Tidskrift utgiven av Riksförbundet för laryngektomerade)

[3] Robbins,J, Fischer,H, Blom,E & Singer,M. (1984) "A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production" J of Speech and Hearing Disorders, Vol 49, pp 202-210, May 1984

[4] Robbins,J , Fischer,H, Blom,E & Singer,M. (1984) "Selected acoustic features of tracheoesophageal, esophageal, and laryngeal speech" Arch Otolaryngol, Vol 110:670-672, Oct 1984

[5] Tardy-Mitzell,S, Andrews.M & Bowman,S (1985)"Acceptability and intelligibility of tracheoesophageal speech" Arch Otolaryngol, Vol 111, pp 213-219, April 1985

[6] Gandour,J & Weinberg,B (1985) "Production of intonation and contrastive stress in esophageal and tracheoesophageal speech" Journal of Phonetics 13, pp 83-95

[7] Weinberg,B, Gandour,J (1986) "Prosody in alaryngeal speech" from 'Current Strategies of Rehabilitation of the Laryngectomized Patient' Seminars in Speech and Language, Vol 7, No 1, pp 95-106, February 1986, Shanks,J (Editor)

[8] Hammarberg,B. & Nord,L. (1987) "Matstrupstal kontra tal via talventil - akustisk analys av två sätt att tala efter laryngektomi." Paper presented at the symposium 'Tal, Ljud, Hörsel', Lund, October 1987

[9] Hammarberg, B. & Nord, L. (1987) "Traditional esophageal speech versus tracheo-esophageal speech - acoustic and perceptual aspects." Paper presented at the 'Seminar on Post Laryngectomy Voice Rehabilitation', Göteborg, November 1987

[10] Hammarberg,B & Nord,L (1988) Phoniatric and Logopedic Progress Report, Huddinge Univ Hospital (forthcoming)

[11] Nord,L & Hammarberg,B (1988) STL-QPSR report from the Department of Speech Communication and Music Acoustics, Royal Institute of Technology (KTH) (forthcoming)

[12] Rothenberg, M (1973) "A new inverse-filtering technique for deriving the glottal airflow waveform during voicing", JASA 53, pp 1632-1645

[13] Hammarberg,B. (1986) "Perceptual and acoustic analysis of dysphonia." Studies in Logopedics and Phoniatrics No 1 (Doct Dissertation), Dept of Logopedics and Phoniatrics, Huddinge Univ Hospital, Huddinge

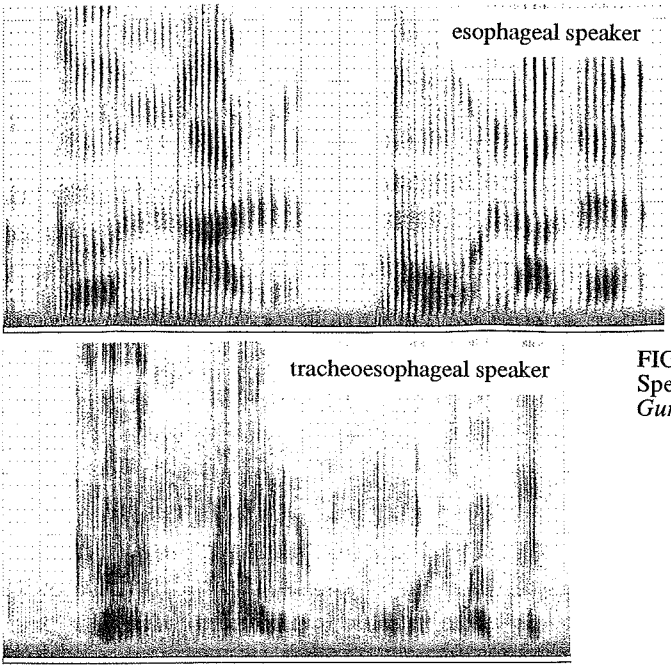

FIGURE 1
Spectrograms of
Gunnar målade...

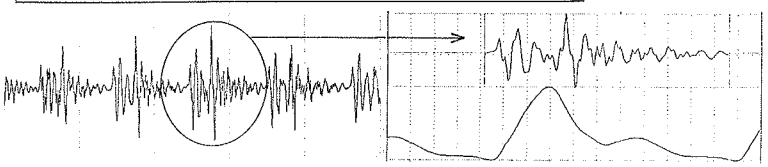

FIGURE 2 Vowel /a/ Extracted voice source pulse (TE-speech)

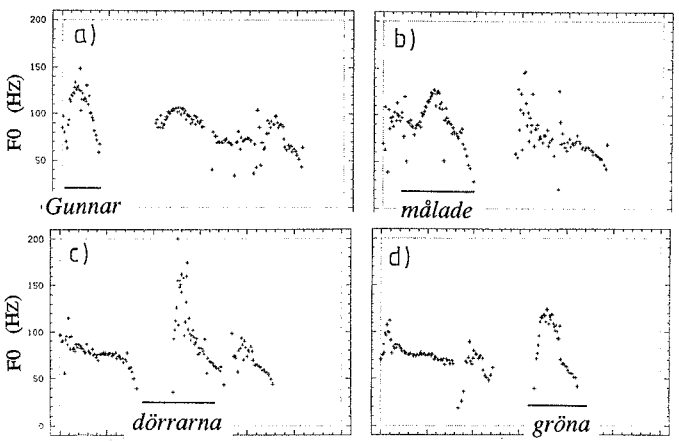

FIGURE 3 F0-tracing of sentences with emphatic stress (TE-speech)