

CONTINUED INVESTIGATIONS OF LARYNGECTOMEE SPEECH IN NOISE - MEASUREMENTS AND INTELLIGIBILITY TESTS

Lennart Nord, Britta Hammarberg*, and Elisabet Lundström

Department of Speech, Music and Hearing, KTH, STOCKHOLM, SWEDEN

*also Dept of Logopedics and Phoniatrics, Karolinska Institute, Huddinge University Hospital

Phone +46 8 790 7874, Fax +46 8 790 7854, e-mail: lennord@speech.kth.se

ABSTRACT

The speech of nine laryngectomized persons is analysed. Both tracheo-esophageal and esophageal speakers are included. The speech performance of the subjects is evaluated while they are reading texts aloud with varying amounts of noise in their ears. One hypothesis is that the amount of noise will influence the articulation skill and voice behaviour of the speakers. Preliminary results show that some of the tracheo-esophageal speakers were able to raise their voice level as much as the normal laryngeal speakers. The esophageal speakers on the other hand were usually not able to produce as strong voice levels during the text readings. Acoustic speech parameters, such as sound pressure and spectral characteristics were measured and compared among the subjects. When the speakers were forced to use a great deal of effort, they spent air rapidly and had to use shorter stretches of speech with many pauses for inhalation.

1. INTRODUCTION

In an ongoing project, the speech of laryngectomized persons is analysed. Two types of alaryngeal speaking techniques are used by the subjects: (1) esophageal speech, E-speech or (2) tracheo-esophageal fistula speech, TE-speech. In both techniques the voice source is the vibratory upper part of the esophagus. The two techniques differ as regards air support for phonation. In E-speech, air is injected into the esophagus from the mouth and in TE-speech the pulmonary air is the driving force through the fistula in the wall between trachea and esophagus. Within the project different aspects of alaryngeal speech have been in focus, e.g. intelligibility, acceptability and acoustic measurements, see e.g. refs [1,2,3]. In previous recording sessions, which were made in an anechoic chamber, we asked the subjects to read with different voice efforts in order to assess the effect of loudness on their voices [4]. However, our results showed that the speakers seldom raised their voice levels, resulting in more or less equal levels for the two speaker groups. This might be an effect of reading in an anechoic chamber, which is artificially silent.

2. PURPOSE

In the present study, we have continued with more subjects and can now present results for nine laryngecto-

-mees. The speakers were asked to read aloud while exposed to noise at different levels in their ears, the hypothesis being that the amount of noise would influence articulation skill and voice behaviour of the speakers.

3. SUBJECTS AND RECORDINGS

The subjects in the present study were 5 tracheo-esophageal (TE) speakers (between 55 and 68 years old) and 4 esophageal (E) speakers (between 52 and 72 years old). All were native Swedish speakers living in the Stockholm area, and recognized as fluent speakers. The TE speakers were using the Blom-Singer low pressure prosthesis [5] and digital occlusion during speech.

The speech performance of the subjects was evaluated while they were reading texts aloud with varying amounts of noise in their ears. The subjects read short texts, taken from everyday prose (H.C. Andersen, A. Henriksson, C. Hylinger and T. Danielsson) under three noise conditions, 0 dB, 70 dB and 80 dB. The noise consisted of a number of voices in a cacophony. The subjects were instructed to raise their voices in order to be heard. During the reading, the absolute sound pressure level at 0.3 meter distance was calculated.

4. LISTENERS

To evaluate the intelligibility in noise background, five normal-hearing listeners were asked to adjust the level of noise while exposed to tape-recorded readings of the laryngectomized speakers. The listener task was to use as high a level of noise as possible, while still being able to comprehend what was said ("just follow conversation") [6]. No separate assessment was made of their actual understanding of the text. Other studies in this area are [7] at our department and [8] by our group.

The listener group (2 females, 3 males) was experienced in participating in listening tests. Ages ranged from 25 to 51 years.

5. ANALYSIS

Calibrated long-term average spectra were produced from the tape-recorded readings by a Hewlett Packard 3562A Dynamic Signal Analyzer. We also compared the level of the fundamental during the reading with the sound pressure level (SPL). The fundamental (F0) was identified as the peak in the low frequency area of the power spectrum displayed on the screen. However, for many of the alaryngeal voices identification of the fundamental turned out to be difficult as these speakers have a very weak and aperiodic fundamental. Further-

more, the recording booth introduced some noise in the low frequency region, at about 70 Hz, obscuring part of the spectrum, where a few of the speakers had their fundamental according to earlier measurements.

6. SPEECH RATE AND PAUSING

As regards speech tempo and pausing, we calculated the number of pauses versus loudness. As the texts were not identical for the speakers, the various text-readings were normalized over a hundred-word paragraph. Our hypothesis was, that when the E-speakers were forced to use a great deal of effort, they spent air rapidly and had to use shorter stretches of speech with many pauses for inhalation.

The same effect could be true also for the TE-speakers, although not to the same extent. In Table I, the number of pauses during the text reading is tabulated together with the SPL for different noise conditions. As can be seen, the speakers increased the number of pauses with increasing SPL, although individual speaking styles appeared. Thus, speaker TE2 paused quite frequently and also used a strong voice level throughout all his readings, regardless of noise level. The normal speaker used fewer pauses and did not increase the number of pauses with increasing voice effort. In a similar way speaker TE1 also used quite few pauses.

7. INTELLIGIBILITY TESTS

The test was quite easy to carry out and the listener performances were stable. Table 3 shows mean values of the relative amount of noise compared to the speaking levels. As can be seen listening to the esophageal speakers allowed the least amount of competing noise, while the tracheo-esophageal speakers allowed more noise. The normal speaker on the other hand was still more intelligible in the sense that listeners allowed even more babble noise on his readings.

8. RESULTS AND DISCUSSION

The SPL results are displayed in Figure 1. The speaking distance is set to 0.3 m. The first tracheo-esophageal speaker (TE1) was able to raise his voice level from 66 to 79 dB with increasing noise in his ears. This is within the range of normal laryngeal speakers. The second TE-speaker (TE2) had a quite strong voice already at normal voice effort and did not change its level (71-72 dB). The first two esophageal speakers on the other hand were not able to produce as strong voice levels during the text readings (58-63 and 65-68 dB, respectively). A limited number of articles reporting on long-term average spectral analysis of the alaryngeal voices have been published. In Weinberg et al. [9], a comparison was made between esophageal speech and normal laryngeal speech and differences in intensity were found to be about 10 dB.

As regards the relation between level of fundamental and SPL, we got some difference values, that point in the direction that when a speaker is raising his voice level,

more energy is attributed to the formant region and relatively speaking less to the fundamental, see Figure 2.

Acoustic speech parameters were measured and compared among the subjects. As regards speech tempo and pausing, we quantified the number of pauses versus loudness. Our hypothesis was, that when the E-speakers were forced to use a great deal of effort, they spent air rapidly and had to use shorter stretches of speech with many pauses for inhalation. The same effect could be true also for the TE-speakers, although not to the same extent. In Table I, the number of pauses during the text reading is tabulated together with the SPL for different noise conditions. As can be seen, the speakers increase the number of pauses with increasing SPL, although individual speaking styles appear. Thus, speaker TE2 is pausing quite frequently and also using a strong voice level throughout all his readings, regardless of noise level.

In Table 2, results from the listening test ("just follow conversation") also revealed a rank order regarding listeners' ease of following conversation in babble noise: 1) normal speaker, 2) tracheo-esophageal speaker, and 3) esophageal speaker. Cf our earlier study [8].

9. CONCLUSION

This type of test method with speech in background noise still seems promising for assessment of voice effort. The TE-speakers seem to be able to raise their voice levels as high as the laryngeal speaker, while the E-speaker usually did not. Also the method with "just follow conversation" seems also useful in the comparison of the quality of the different speaking techniques.

10. ACKNOWLEDGEMENTS

The study was supported by grants from the Swedish Council for Social Research (project No. C90-0172:2) and the Swedish Cancer Society (project No 3000-B91-02XAA). We are grateful to Nick Edsberg, Dept of Radiology, Huddinge Univ. Hospital, for making the videofluoroscopic registrations, and to Per-Åke Lindestad, Dept of Logopedics and Phoniatics, Huddinge Univ. Hospital, for the fiberoendoscopy.

11. REFERENCES

- [1] B. Hammarberg, and L. Nord, "Communicative aspects of laryngectomy speech. Presentation of a project and some preliminary results". Phoniatic and Logopedic Progress Report No 6, pp. 10-27, Dept of Logopedics and Phoniatics, Huddinge Univ Hospital (Karolinska Institutet), Sweden, 1988
- [2] B. Hammarberg, E. Lundström and L. Nord, "Consonant intelligibility in esophageal and tracheo-esophageal speech. A progress report". Phoniatic and Logopedic Progress Report No 7, pp. 49-57, Dept of Logopedics and Phoniatics, Huddinge Univ Hospital (Karolinska Institutet), Sweden, 1990
- [3] L. Nord, L., B. Hammarberg, and E. Lundström, "Voiced-voiceless distinction in alaryngeal speech - acoustic and articulatory observations". Speech

Transmission Laboratory Quart. Progress & Status Report, Dept of Speech Comm & Music Acoustics., STL-QPSR 2-3/1992, pp. 19-22, (KTH), Stockholm, 1992

- [4] B. Hammarberg, E. Lundström, and L. Nord, "Intelligibility and acceptability of esophageal speech, related to acoustic characteristics". Proceedings of the 22nd Congress of Intern. Ass. of Logopedics and Phoniatics, Hannover, Aug 1992.
- [5] A. Perry, Surgical voice restoration following laryngectomy: the tracheo-oesophageal fistula technique (Singer-Blom). British J of Disorders of Communication, vol. 23, pp. 23-30, 1988

- [6] S. Hygge, J. Rönnberg, B. Larsby, S. Arlinger, Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds. J Speech Hear Research 35: pp. 208-215. 1992
- [7] G. Öhngren, M. Dahlquist, Comparisons between different measurements of the auditory, visual, and audiovisual speech perception ability of hearing impaired subjects; some preliminary results. Speech Transmission Laboratory Quart. Progress Status Report, Dept of Speech Comm & Music Acoustics, (KTH), Stockholm, STL-QPSR 4/1988, pp. 23-30, 1988
- [8] L. Nord, B. Hammarberg, and E. Lundström, "Laryngectomee speech in noise - voice effort and intelligibility", EUROSPEECH '93, Berlin, Sept 21-23, 1993, Vol. 3, pp. 1981-1984
- [9] B. Weinberg, Y. Horii, and B. Smith, "Long-term spectral and intensity characteristics of esophageal speech". J Acoust. Soc. Am. 67, pp. 1781-1784, 1980

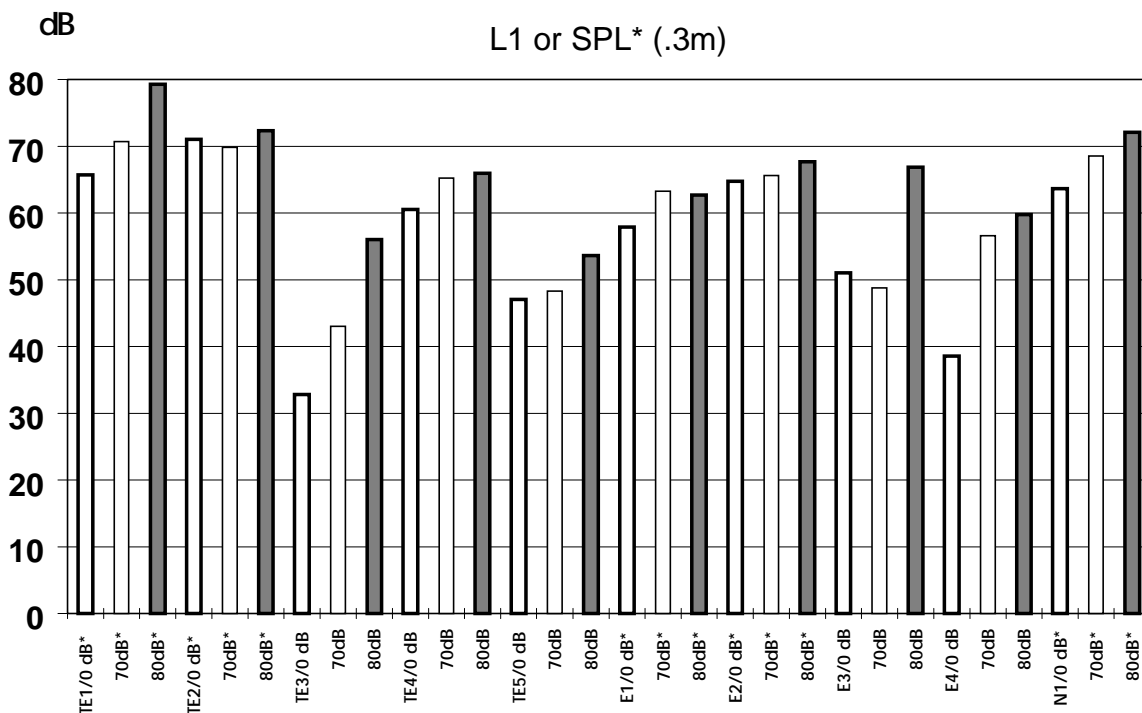


Figure 1. The level of the first formant or the sound pressure level (SPL) for the 10 speakers, exposed to various amounts of noise in their ears.

Table 1. Number of pauses during the speakers' text-readings ((normalized over a hundred-word paragraph) and the sound pressure level (dB, 0.3 m), for different noise level exposures Two esophageal, two tracheo-esophageal and one normally speaking man.

Noise-exposure	E1		E2		TE1		TE2		N1	
	pauses	dB	pauses	dB	pauses	dB	pauses	dB	pauses	dB
0 dB	18	58	18	65	13	66	19	71	08	74
70 dB	29	63	24	66	15	71	17	70	10	79
80 dB	26	63	22	68	16	79	23	72	11	82

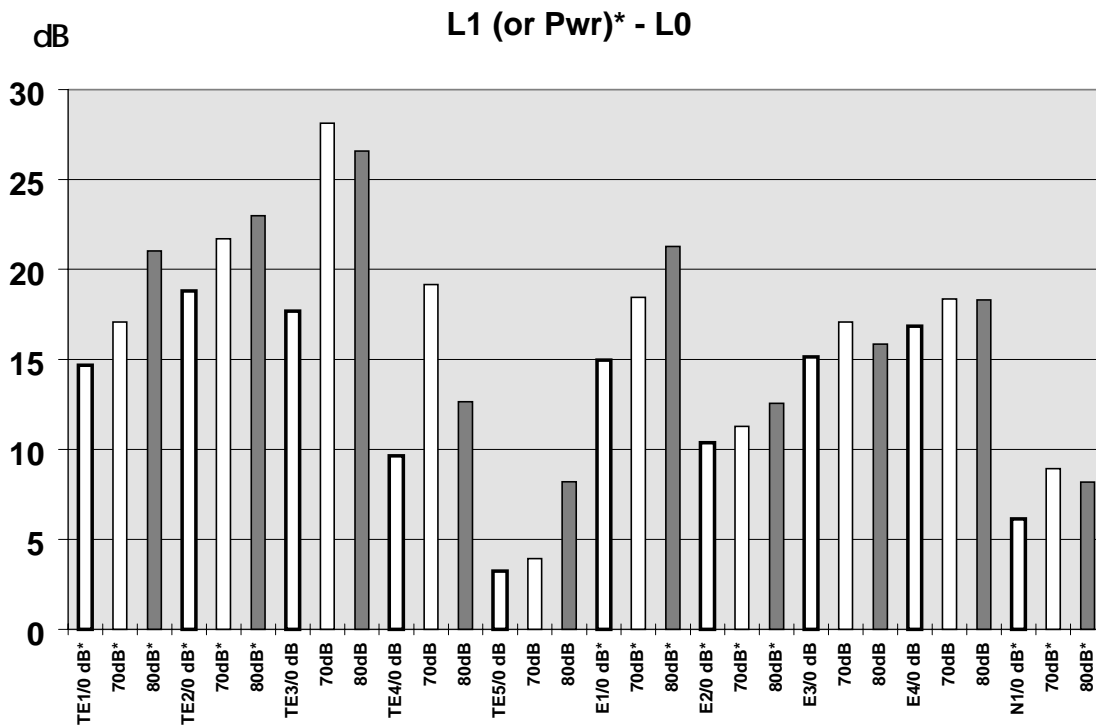


Figure 2. Difference between the level of the first formant (L1) or the sound pressure level (SPL) and the level of the fundamental (L0) for the 10 speakers, exposed to various amounts of noise in their ears.

Table 2. Part of the results of the intelligibility test, 5 listeners. “Just follow conversation” (=JFC-test): Difference level between noise level and speaking level

	mean	stdev		mean	stdev		mean	stdev
E1/0 dB	1,12	3,54	TE1/0 dB	3,90	1,14	N1/0 dB	3,56	1,30
E1/70 dB	0,94	2,68	TE1/70 dB	3,72	1,67	N1/70 dB	6,66	2,39
E1/80 dB	2,11	3,70	TE1/80 dB	1,13	2,30	N1/80 dB	10,13	2,86
E2/0 dB	1,64	1,95	TE2/0 dB	2,60	2,97			
E2/70 dB	1,20	3,35	TE2/70 dB	3,78	2,97			
E2/80 dB	1,32	1,87	TE2/80 dB	3,88	2,77			