Institute of Phonetic Sciences,
University of Amsterdam,
Proceedings 20 (1996), 53 - 64.

INTELLIGIBILITY OF ALARYNGEAL SPEECH
A DESIGN FOR RESEARCH

*Corina J. van As, Florien J. Koopmans-van Beinum and Frans J.M. Hilgers**

Abstract

In a joint project of the University of Amsterdam (Institute of Phonetic Sciences), the University of Nijmegen (Speech- and Language Pathology), and the Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital (Department of Otolaryngology-Head and Neck Surgery), alaryngeal speech is investigated.
All patients in this study use tracheoesophageal (TE) speech. To obtain TE speech, a fistula is created between the trachea and the oesophagus; this opening allows insertion of a voice prosthesis. The prosthesis acts as a one-way valve through which pulmonary air can be directed into the oesophagus, by means of which the pharyngoesophageal (PE) segment starts vibrating. The patient has to close his stoma to be able to speak.
The goal of the project is to obtain insight in the acoustical, perceptual, and clinical aspects of alaryngeal voice quality, and intelligibility.
In the present paper, first surgical and anatomical aspects of total laryngectomy, and methods of voice restoration are described. Then the speaker groups participating in the project are discussed, and the type of data that will be gathered from these speakers are described. In order to decide what recordings have to be made to measure intelligibility, the project started with a study of the literature on intelligibility. Conclusions are drawn for the design of the project.

1. Introduction

1.1. Total laryngectomy

Laryngeal carcinoma is the most frequent head and neck malignancy in the Netherlands. Approximately 700 new cases per year are detected (de Winter et al., 1989). Fortunately, most patients present themselves with early disease and can still be treated by curative radiotherapy with preservation of all laryngeal functions, especially the voice. Only advanced cases or recurrences after radiotherapy have to be treated with surgery. In most cases this means that a total laryngectomy has to be performed. In the Netherlands about 250 patients per year are laryngectomized (Ackerstaff et al., 1994). Due to the removal of the entire larynx, the trachea has to be sutured to the skin in the base of the neck to form a tracheostoma. The result is a permanent disconnection between the airway and the alimentary tract and between the upper (c.q. nose) and lower airways.

* The Netherlands Cancer Institute (Antoni van Leeuwenhoek Hospital), Amsterdam.

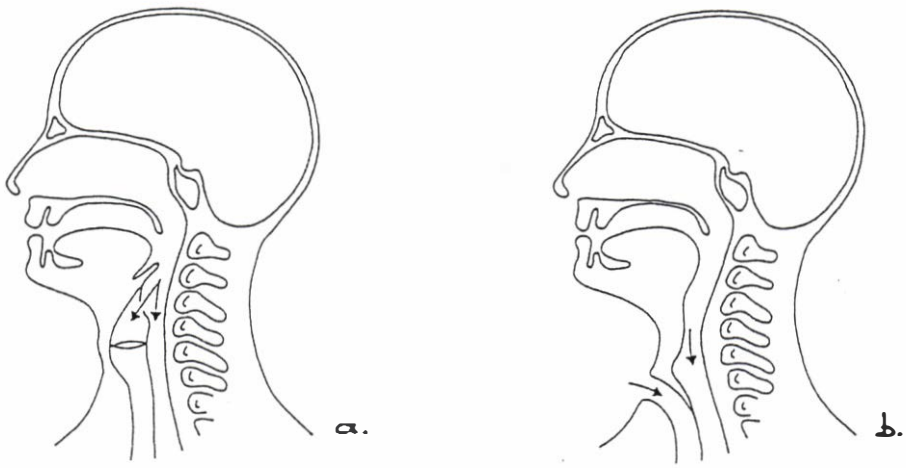

Figure 1a. Anatomical situation before surgery. 1b. Anatomical situation after surgery.

When the patient is recovered from the operation, he can eat through the normal route, but he is not able to speak in the normal way, because of the loss of the normal voicing source and the disconnection between the lower airway and oral/nasal cavity. Breathing takes place via the tracheostoma in the neck. In Figure 1 schematical situations before and after surgery are shown.

1.2. Voice rehabilitation

A total removal of the larynx and thus of the vocal cords, implies the loss of the natural voicing source. Already since the first laryngectomy by Billroth in 1873 described by Gussenbauer (1874), voice restoration was a very important factor. Until 1970 the two most widely used voice rehabilitation methods were training in esophageal (E) speech, or the use of an electrolarynx. In esophageal speech the patient is trained to insufflate air from the mouth into the esophagus, which makes the upper sphincter (pharyngoesophageal (PE) segment) of the esophagus vibrate. An electrolarynx emits pulses through the pharyngeal wall into the oral cavity.

Since 17 years, tracheoesophageal (TE) speech is used as a method for voice restoration. Consistently high success rates have been reported after the first description of a useful prosthetic device by Singer & Blom (1980). For TE speech a fistula is created between the trachea and the esophagus; this opening allows insertion of a prosthesis, which acts as a one-way valve through which pulmonary air can be directed into the esophagus. The patient has to close his tracheostoma to be able to speak. Figure 2 shows the schematical situation in patients using tracheoesophageal speech.

The main difference between E speech and TE speech is the air supply, which is 60-80 cc for E speech, and several liters in the pulmonary driven TE speech. With TE speech, compared to E speech, a higher number of patients achieve an acceptable voice (Hilgers & Balm, 1993).

Athough different studies have shown the advantages of tracheoesophageal speech above esophageal speech, speech related problems are frequently noted in literature (Natvig, 1984; Jay et al., 1991; Jones et al., 1992; Dhillon et al., 1982). Natvig (1984) stated that, for 40% of the 186 patients he interviewed, loss of the normal voice was the greatest problem, independent of their postoperative speech intelligibility.

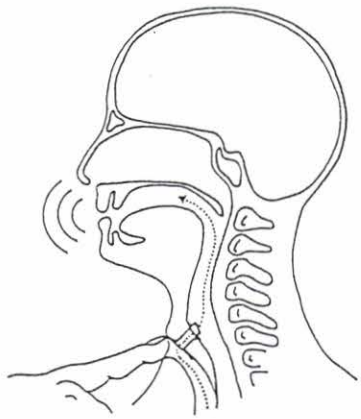

Figure 2. Anatomical situation in tracheoesophageal speech

2. Design of the project

2.1. Purpose of the project

The main aim of the present project is to develop a method to evaluate the speech of laryngectomized patients, and to obtain more insight in alaryngeal speech properties, and the vibration pattern of the PE segment. This will be done by the investigation of perceptual, acoustical, and clinical aspects of voice quality and intelligibility, and by investigation of relations between those aspects.

By means of the evaluation method to be developed, we want to investigate the influence of the Provox voice prosthesis that the patients use, the use of a stomafilter, and the age of the speakers. Furthermore we want to obtain more insight in changes in alaryngeal voice quality during the first year after surgery.

2.2. Speaker groups

All speakers are patients treated at the Netherlands Cancer Institute. One group is recorded once, and consists of 40 TE speakers, who had surgery over 6 months ago. A longitudinal group consists of about 20 TE speakers, that are laryngectomized during the first year of the study; recordings are made at 1-2 months, 5-7 months, and 11-13 months after surgery. Apart from these groups, also an age-, and sex-matched control group of 20 speakers will be recorded once.

2.3. Data collection

The following data will be gathered in each recording session:

- 3 times a sustained /a/ at comfortable pitch and loudness;
- read-aloud text (about 3 minutes);
- phonetogram;
- videostroboscopy or fiberoptic recordings of the PE segment;
- electroglottographic (EGG) assessment of the voice source;
- videofluoroscopic recordings of the PE segment;

- aerodynamic tests;
- questionnaires;
- intelligibility tests.

In order to be able to make a decision about the exact speech material that has to be gathered for the acoustical analyses, for the perceptual evaluations, and for the recordings of the PE segment, first the intelligibility test has to be performed. The consonants that are difficult to perceive correctly, can than be studied more extensively.

3. Studies on Intelligibility

In studies on intelligibility of alaryngeal speech, as reported in literature, various speech stimuli, listener groups, and research methods have been used. Intelligibility of laryngectomees is investigated in (American) English, Swedish, Spanish, German, and Dutch. The speech material, speakers, and methods chosen, depend upon the purpose of the study. The literature study described below, is performed to come to a descision about what research method should be used, which listeners should participate, and what speech material has to be gathered for the research project. In the present project we are especially interested in the influence of the new voicing source (PE segment) on intelligibility of consonants produced by TE speakers. First we want to perform an intelligibility test, to establish which Dutch consonants are difficult to perceive for listeners. For the so found "problem consonants" the vibration characteristics of the PE segment will be extensively investigated by videofluoroscopic and videostroboscopic recordings of the PE segment.

3.1. Aims of research reported in literature

Intelligibility of alaryngeal is investigated for different reasons:

(1) to compare different alaryngeal speech techniques with each other and with normal speech. Comparisons are made for normal speech, esophageal speech, tracheoesophageal speech, and speech with use of an electrolarynx;
(2) to compare different voice prostheses and stoma occlusion conditions in tracheoesophageal speech;
(3) to investigate consonant intelligibility in laryngectomees;
(4) to investigate vowel intelligibility in laryngectomees.

3.2. Speech material

As speech material in the reported studies, different stimuli are used, varying from conversational speech to nonsense syllables. The material used consists of conversational speech, uninterrupted discourse, read-aloud text, sentences of different lengths, questions, bisyllabic words, monosyllabic words, and nonsense words. The words are sometimes presented in a carrier phrase. Speech material used in the different studies is given in Table 1.

Table 1. Speech stimuli used in studies on intelligibility of alaryngeal speech. Used speech material was conversational speech (CS), uninterrupted discourse (UD), reading text aloud (RA), sentences of different lengths (S), questions (Q), bisyllabic words (W2), monosyllabic words (W1), and nonsenswords (N).

Study	CS	UD	RA	S	Q	W2	W1	N
Christensen & Weinberg (1976)								x
Clark & Stemple (1982)				x				
Cullinan et al. (1986)		x	x					
De Maddalena & Zenner (1996)			x				x	
Doyle et al. (1988)								x
Doyle et al. (1990)								x
Fujimoto et al. (1991)	x			x			x	
Hammarberg et al. (1990)					x		x	
Hammarberg et al. (1992)							x	
Hammarberg & Nord (1995)							x	
Hubbard & Kushner (1980)			x					
Kalb & Carpenter (1981)							x	
Mahieu et al. (1986)							x	
Man-Lai Yiu et al. (1994)							x	
Miralles & Cervera (1995)						x		
Nichols (1976, I & II)							x	
Nord et al. (1992)							x	
Nord et al. (1993)			x					
Pindzola & Cain (1988)							x	
Rizer et al. (1984)							x	
Robbins et al. (1986)								x
Williams et al. (1989)	x	x					x	
Williams et al. (1990)	x	x	x	x			x	

Table 2. Research methods used in investigating intelligibility of alaryngeal speech. Used research methods are listeners judgements on a 5- or 7-point scale (SC), answering questions (AQ), sentences (TS), words (TW), or phonemes (TP), multiple choice/rhyming words (MC), spectrographic analysis (SA), allowed noise level (NL), and telephone test (TT).

Study	SC	AQ	TS	TW	TP	SA	NL	MC	TT
Christensen & Weinberg (1976)					x				
Clark & Stemple (1982)							x		
Cullinan et al. (1986)	x								
De Maddalena & Zenner (1996)			x	x					x
Doyle et al. (1988)					x				
Doyle et al. (1990)					x				
Fujimoto et al. (1991)	x		x	x					
Hammarberg et al. (1990)		x		x					
Hammarberg et al. (1992)				x					
Hammarberg & Nord (1995)				x					
Hubbard & Kushner (1980)			x						
Kalb & Carpenter (1981)				x					
Mabieu et al. (1986)							x		
Man-Lai Yiu et al. (1994)				x					
Miralles & Cervera (1995)				x					
Nichols (1976, I & II)				x					
Nord et al. (1992)				x					
Nord et al. (1993)							x		
Pindzola & Cain (1988)								x	
Rizer et al. (1984)								x	
Robbins et al. (1986)								x	
Williams et al. (1989)	x								
Williams et al. (1990)	x								

3.3. Patients and methods

Speakers

All speakers were laryngectomized patients using an electrolarynx, or producing esophageal, or tracheoesophageal speech. In few studies speakers who are able to use both, esophageal, and tracheoesophageal speech, participate.

Listeners
In the various studies, different listener groups are participating. Their level of experience, regarding listening experiments and/or alaryngeal speech varied from experienced to inexperienced. In most studies, no differences are found between judgements of different groups. Naive listeners are chosen more often, because that is the listener group a laryngectomee has to deal with in every day situations.

Methods
In studying intelligibility of laryngectomees, several research methods are used. A global judgement of intelligibility in conversational speech, uninterrupted discourse, and sentences, is performed using a 5- or 7-point scale, on which the overall intelligibility is rated. In more precise methods listeners transcribe what they heard, these transcriptions can be used subsequently to calculate a percentage of correctly understood sentences, words, or phonemes. To study the acoustical characteristics of the consonants, also spectrographic analysis is used. In a few studies the intelligibility of alaryngeal speech in noise is investigated, listeners then have to adjust the noise level to a level on which they can just understand what de patient is saying. In a study performed in Germany, intelligibility is studied via telephone. The various methods are summarized in Table 2.

Statistics
The following statistical analysis methods are reported:
(1) calculating the mean overall intelligibility scores, on the basis of perceptual scales;
(2) calculating the level of allowed noise;
(3) calculating the percentage of correctly understood sentences, words, or consonants;
(4) registration of errors in confusion matrices;
(5) feature analysis on the basis of pooled confusion matrices, or SINDSCAL (Symmetric Individual Differences Scaling) analysis.

3.4. Results of studies reported on in literature

In our own research project, we are interested primarily in listeners' perceptions of Dutch consonants in tracheoesophageal speech. Therefore, only results of earlier studies investigating intelligibility of alaryngeal speakers at a consonantal level are discussed.

Hammarberg et al. (1990)

Consonant intelligibility was investigated using four lists of 30 monosyllabic, Swedish words (CVC). Audio recordings of these words, read by four esophageal (E) and three tracheoesophageal (TE) speakers, were judged by seven listeners who had no previous experience with alaryngeal speech. When mispronouncations were noted, the speaker was given an extra chance by reading the word aloud, and inviting the speaker to repeat it. The listeners were asked to write down the word presented just as they heard it, using standard orthography. The mean consonant intelligibility was calculated, the percentage identification errors was computed for both initial and final consonants, and a confusion matrix was made for the initial consonants. In the TE speaker group, the mean consonant intelligibility was 87%, the percentage identification errors was 11% for the initial consonants, and 4% for the final consonants. For

the initial consonants it appeared from the confusion matrix, that a /p/ was perceived as a /b/ in 30%, a /b/ was never perceived as /p/, a /d/ was perceived as /t/ in 10%, a /t/ as /d/ in 3% of the cases. The /g/ was perceived as /k/ in 57% of the cases. For nasals an /n/ was perceived as /m/ in 30%, an /m/ as /n/ in 9% of the cases. The liquids /r/ and /l/ were perceived correct in almost all cases. Except for the /h/, the fricative sounds were not difficult for the listeners to perceive correctly. Omissions of /h/ were noted in 67% of the cases.

Hammarberg et al. (1992)

This study was the same as the study by Hammarberg et al. (1990) described above, only more speakers participated. The word list was read by 5 esophageal speakers, and 4 tracheoesophageal speakers. The results found in the first study, still remain the same.

Nord et al. (1992)

This study refers to the two earlier studies of Hammerberg et al. (1990, 1992). Two spectrographic illustrations shown in this paper support the earlier findings. Spectrograms of the word pairs /bank/ and /pank/, and /gal/ and /kal/ are shown. Based on these illustrations they state that ..."A typical Swedish initial /p/ usually has a clear burst interval and an F1 cut-back. In this production of /p/ it is difficult to see any trace of burst and F1 starts quite sudden. The word /pank/ is thus perceived as /bank/ with reduced voicing cue but still not an initial /p/." (p. 20), and ..."which was meant to be /gal/ but was perceived as /kal/. In the spectrogram there is no initial voicing and an initial very irregular voiced segment is heard as a burst interval, thus signalling a /k/." (p. 20).

Doyle et al. (1988)

Consonant intelligibility is investigated in both TE and E speech. The speech material consisted of CVCVC nonsense stimuli, in which the initial and final C always was an /m/, the V was an /i/ or /u/, and the medial C was one of the twenty-four English consonants. Each stimulus was embedded in the carrier phrase "Say .....now". Each speaker read 24 pratice items, before the real recordings were made. Listeners were 15 normal hearing people, with no prior experience with alaryngeal speech. They were able to use the International Phonetic Alphabet (IPA). Listeners' responses were converted to confusion matrices, and analyzed for overall intelligibility, voicing and manner features, and consonant omissions. The mean overall intelligibility score for the TE speakers was 65%. Voiced consonants were correct in 60%, voiceless consonants in 20%. Liquid-glides produced by TE speakers were best received, followed by nasals, plosives, fricatives, and affricates. In speakers perceived as the most intelligible, listeners' result showed that they had the fewest omissions.

Doyle et al. (1990)

Speech material, speakers, and methods were all the same as in the study of Nord et al.(1992) described before, but another analysis method is used. The listeners' perceptual responses were analyzed using Symmetric Individual Differences Scaling

(SINDSCAL) to determine whether distinctive feature differences existed between esophageal and tracheoesophageal speech.

Although greater perceptual weightings were observed for TE speech, these productive/perceptual features were weighted similarly for both speech methods. In the analysis the lack of a separate voicing dimension is noted. This is related to the difficulty that alaryngeal speakers have in producing the voicing distinction.

Nichols (1976)

Consonant intelligibility was investigated in 11 English esophageal (E) speakers. Three different listener groups, with different levels of experience, participated in this study. Speech stimuli were monosyllabic (CVC) words. From each word both consonants were used for analysis. The words were embedded in the carrier phrase "Say ..... again". Listeners were asked to write down the CVC word. The overall word intelligibility was 36%. The /m/ was the most intelligible with 69% correct, the /p/ was the least intelligible, with 15% correct. The mean consonant intelligibility was 53%.

Miralles & Cervera (1995)

Consonant intelligibility was studied in 30 alaryngeal Spanish speakers: 20 tracheoesophageal, and 10 esophageal talkers were recorded reading a list of 24 2-syllabic Spanish words. Speech stimuli were presented to a group of 140 listeners using an open response paradigm. Errors were registered in confusion matrices, and analyzed in terms of phonetic dimensions. In the TE talkers, the confusions included /p/ with /b/ (12.9%), /b/ with /p/ (8.7%), /d/ with /t/ (7.5%), and /g/ with /k/ (8%). The authors state that ..."These types of confusions suggest that the voice-voiceless distinction is difficult to perceive, when the speaker is a laryngectomized patient. Another portion of the errors could be attributed to the pathology itself, such as the difficulty in producing velars, as in both groups of laryngectomized patients."(p. 567).

3.5. Discussion

Although all these studies are especially on consonant intelligibility, the used speech material, patients, and methods are different. Speech material consists of two-syllabic words (Miralles & Cervera, 1995), one-syllabic words (Hammarberg et al., 1990, 1992; Nord et al., 1992; and Nichols, 1976), or nonsens words (Doyle et al., 1988, 1990). Nonsense stimuli are used to avoid effects of linguistic cues (phonological and semantic) inherent in meaningful material (Doyle et al., 1988, 1990). The words and nonsense stimuli are sometimes embedded in a carrier phrase (Doyle et al., 1988, 1990), Nichols, 1976). From both these studies it becomes not clear why a carrier phrase is used. Hammarberg et al. (1990, 1992), Nord et al. (1992), and Nichols (1976) studied initial and final consonants. Doyle et al. (1988, 1990) studied medial consonants, and Miralles & Cervera (1995) studied initial and medial consonants. In almost all studies the listeners wrote down what they heard, using standard orthography, only in research of Doyle et al. (1988, 1990), listeners wrote their responses in the International Phonetic Alphabet (IPA). Data were analyzed using confusion matrices (Hammarberg et al., 1990, 1992; Doyle et al., 1988; Nichols, 1976; and Miralles & Cervera, 1995), feature analysis (Doyle et al., 1988, 1990), or spectrographic analysis (Nord et al., 1992).

In all studies, the participating listeners do not have any experience with laryngectomized speakers, though some speaker groups do have experience with listening experiments.

For the purpose of our research project, consonant intelligibility will be investigated in all three positions, i.e. initial, medial, and final, because it is possible that the production of a consonant is influenced by its position in a word. Since it becomes not clear why some studies do use carrier phrases, and some studies do not, a decision has to be made on the basis of the purpose of the present project. The interest is on consonant intelligibility, thus we want those consonants not to be influenced by any surrounding words or phonemes. Due to coarticulation effects it seems not wise to let the speaker use a carrier phrase, possible influences of surrounding phonemes are so avoided; when presenting the stimulus to the listener a carrier phrase can be added to the stimulus, just to alert the listener. In testing phoneme intelligibility in speech synthesis systems, it is prefered to use nonsense words and an open response task, because this provides excellent diagnostic material (Pols & SAM partners, 1992). By using nonsense stimuli, it is easy to obtain the same amount of data for each consonant, even for those consonants not occurring very often in Dutch, making a reliable analysis possible. Another advantage of using nonsense stimuli is that the researcher is able to structure the consonant-vowel combinations. Therefore, in this study, it seems best to use nonsense stimuli of CV, VCV, and VC formation which are also used in the earlier mentioned SAM project (Pols & SAM partners, 1992). In these CV, VCV, and VC combinations, all single consonants of Dutch are tested in initial, medial, and final position, in combination with a close-front (/i/), a close-back (/u/), and an open vowel (/a/). These vowels represent the three most extreme articulation positions in vowels, by combining each consonant with all three vowels, a possible influence of adjacent vowels can be studied.

Recordings will be made of 15 laryngectomized Dutch speakers who are all using tracheoesophageal speech, these recordings will be judged by 20 naive listeners, with no prior experience to laryngectomees and their speech. The choice is on naive listeners, because this is the group of listeners in which the laryngectomee has the most social contacts.

4. Conclusion

It can thus be concluded that in the part of our project where we intend to investigate consonant intelligibility, the speech material will consist of nonsense stimuli, of CV, VC, and VCV combinations, with the C being all possible Dutch consonants in that position, and the V being /a/, /i/, and /u/. Speakers are 15 laryngectomized male patients. Listeners are 20 naive listeners. The results will be registered in a confusion matrix, and a feature analysis will be performed. On the basis of the results of the study on consonant intelligibility in laryngectomized speakers, it should be possible to pass judgement about which consonants or features cause difficulties for laryngectomees to pronounce. The consonants that appear to be 'misheared' often by the listeners, need further study in the whole project. Thus the data that will be gathered for the whole project are dependant on the outcome of the intelligibility study that will be performed first.

References

Ackerstaff, A.H., Hilgers, F.J.M., Aaronson, N.K., Balm, A.J.M. (1994). "Communication, functional disorders and life style changes after total laryngectomy". *Clin. Otolaryngol.* **19**: 295-300.

Christensen, J.M. & Weinberg, B. (1976). "Vowel duration characteristics of esophageal speech". *J. Speech. Hear. Res.* **19**: 678-689.

Clark, J.G. & Stemple, J.C. (1982). "Assessment of three modes of alaryngeal speech with a synthetic sentence identification (SSI) task in varying message-to-competition ratios". *J. Speech. Hear. Res.* **25**: 333-338.

Cullinan, W.l., Brown, C.S., Blalock, P.D. (1986). "Ratings of intelligibility of esophageal and tracheoesophageal speech". *J. Comm. Disord.* **19**: 185-195.

De Maddalena, H. & Zenner, H.P. (1995). "Evaluation of speech intelligibility after prosthetic voice restoration by a standardized telephone test". *Proceedings 6th international congress on surgical and prosthetic voice restoration after total laryngectomy*. San Sebastian, Spain, 29 sept. - 1 Oct. 1995. Excerpta Medica International Congress Series **1112**: 183-187.

de Winter, G.A., Coebergh, J.W.W., van Leeuwen, F.E., Schouten, L.J. (Eds) (1992). "Incidence of cancer in the Netherlands 1989: first report of the Netherlands Cancer Registry". Utrecht: *Netherlands Cancer Registry-LOK*: 115-139.

Dhillon, R.S., Palmer, B.V., Pittam, M.R., Shaw, H.J. (1982). "Rehabilitation after major head and neck surgery - the patients' view". *Clin. Otolaryngol.* 7: 319-324.

Doyle, P.C., Danhauer, J.L., Reed, C.G. (1988). "Listeners' perceptions of consonants produced by esophageal and tracheoesophageal talkers". *J. Speech. Hear. Disord.* **53**: 400-407.

Doyle, P.C., Danhauer, J.L., Mendel, L.L. (1990). "A SINDSCAL analysis of perceptual features for consonants produced by esophageal and tracheoesophageal talkers". *J. Speech. Hear. Disord.* **55**: 756-760.

Fujimoto, P.A., Madison, C.L., Larrigan, L.B. (1991). "The effects of a tracheostoma valve on the intelligibility and quality of tracheoesophageal speech". *J. Speech. Hear. Res.* **34**: 33-36.

Gussenbauer, C. (1874). "Ueber die erste durch Th. Billroth am Menschen ausgeführte Kehlkopf-Extirpation und die Anwendung eines künstlichen Kehlkopfes". *Arch. Klin. Chir.* **17**: 343-356.

Hammarberg, B., Lundström, E., Nord, L. (1990). "Consonant intelligibility in esophageal and tracheoesophageal speech. A progress report". *Phoniatric and Logopedic Progress report*: 7. Dept. of Logopedics and Phoniatrics, Huddinge University Hospital (Karolinska Institutet), Sweden.

Hammarberg, B., Lundström, E., Nord, L. (1992). "Intelligibility and acceptability of esophageal speech, related to acoustic characteristics". *Proceedings of the 22nd Congress of Inter. Ass. of Logopedics and Phoniatrics*, Hannover, Germany.

Hammarberg, B. & Nord, L. (1995). "The alaryngeal voice source as analysed by videofluoroscopy, fiberendoscopy, and perceptual-acoustic assessment". *Proceedings ICPhS95*, Stockholm, Sweden. Vol. 4: 590-593.

Hilgers, F.J.M. & Balm, A.J.M. (1993) "Long term results of vocal rehabilitation after total laryngectomy with the low resistance, indwelling Provox® voice prosthesis system". *Clin. Otolaryngol.* **18**: 517-523.

Hubbard, D.J. & Kushner, D. (1980). "A comparison of speech intelligibility between esophageal and normal speakers via three modes of presentation". *J. Speech. Hear. Res.* **23**: 909-916.

Jay, S., Ruddy, J., Cullen, R.J. (1991) "Laryngectomy: the patients' view". *J. Laryngol. Otol.* **105**: 934-938.

Jones, E., Lund, V.J., Joward, D.J., Greenberg, M.P., McCarthy, M. (1992). "Quality of life of patients treated surgically for head and neck cancer". *J. Laryngol. Otol.* **106**: 238-242.

Kalb, M.B. & Carpenter, M.A. (1981). "Individual speaker influence on relative intelligibility of esophageal and artificial larynx speech". *J. Speech. Hear. Disord.* **46**: 77-80.

Mahieu, H.F., Schutte, H.K., Annyas, A.A. (1986). "Intelligibility, vocal intensity, and long-term average spectra of Groningen button esophageal speech". In: I.F. Herman (Ed.), *Speech restoration via voice prostheses.* Heidelberg.

Man-Lai Yiu, E., van Hasselt, C.A., Williams, S.R., Woo, J.K.S. (1994). "Speech intelligibility in tone language (Chinese) laryngectomy speakers". *Eur. J. Disord. of Comm.* **29**: 339-347.

Miralles, J.L. & Cervera, T. (1995). "Voice intelligibility in patients who have undergone laryngectomees". *J. Speech. Hear. Res.* **38**: 564-571.

Natvig, K. (1984). "Study no. 5: Problems of everyday life". *J. Otolaryngol.* **13**: 15-22.

Nichols, A.C. (1976). "Confusions in recognizing phonemes spoken by esophageal speakers: I. Initial consonants and clusters". *J. Comm. Disord.* **9**: 27-41.

Nichols, A.C. (1976). "Confusions in recognizing phonemes by esophageal speakers: II. Vowels and diphthongs". *J. Comm. Disord.* **9**: 247-260.

Nord, L., Hammarberg, B., Lundström, E. (1992). "Voice-voiceless distinction in alaryngeal speech - acoustic and articulatory observations". *STL-QPSR* **2-3**: 19-22. Dept. of Speech Comm. & Music Acoustics (KTH), Stockholm, Sweden.

Nord, L., Hammarberg, B., Lundström. E. (1993). "Laryngectomy speech in noise - voice effort, speech rate, and intelligibility". *STL-QPSR* **2-3**: 57-64, Dept. of Speech Comm. & Music Acoustics (KTH), Stockholm, Sweden.

Pindzola, R.H. & Cain, B.H. (1988). "Acceptability ratings of tracheoesophageal speech". *Laryngoscope* **98**: 394-397.

Pols, L.C.W. & SAM-partners (1992). "Multi-lingual synthesis evaluation methods". *Proceedings of ICSLP 92*, Banff, Vol 1: 181-184.

Rizer, F.M., Schechter, G.L., Coleman, R.F. (1984). "Voice quality and intelligibility characteristics of the reconstructed larynx and pseudolarynx". *Otolaryngol. Head. Neck. Surg.* **92**: 635-638.

Robbins, J., Christensen, J., Kempster, G. (1986). "Characteristics of speech production after tracheoesophageal puncture: voice onset time and vowel duration". *J. Speech. Hear. Res.* **29**: 499-504.

Williams, S.E., Scanio, T.S., Ritterman, S.I. (1989). "Temporal and perceptual characteristics of tracheoesophageal voice". *Laryngoscope* **99**: 846-850.

Williams, S.E., Scanio, T.S., Ritterman, S.I. (1990). "Perceptual characteristics of tracheoesophageal voice produced using four prosthetic/occlusion combinations". *Laryngoscope* **100**: 290-293.