VOICE QUALITY JUDGEMENTS AND PHYSIOLOGICAL MEASUREMENTS IN ESOPHAGEAL SPEAKERS WITH AND WITHOUT A GRONINGEN BUTTON.

G.L.J. Nieboer

ENT-Clinic, Voice Res. Lab & Centre for Voice, Sp. and Lang. Dis., Univ. Hosp., and Inst. of Phon. Sc., State Univ, Groningen, The Netherlands.

H.K. Schutte, M.D., Ph.D.

ENT-Clinic, Voice Res. Lab & Centre for Voice, Sp. and Lang. Dis., Univ. Hospital, State University, Groningen The Netherlands.

T. de Graaf, Ph.D.

Institute of Phonetic Sciences, State University, Groningen, The Netherlands.

ABSTRACT.

Physiological measurements have been performed on 26 esophageal speakers, both with and without a Groningen Button. The measured variables are: intra-tracheal, sub- and supra- pseudoglottic pressure, trans-pseudoglottic flow and sound pressure level. Of the same set of speakers, tape recordings were made in view of a perceptual evaluation by a group of 85 judges. The evaluations were done on 13 bipolar semantic scales. The results of both parallel experiments are presented in this contribution, as well as the first results of correlation computations.

1) INTRODUCTION.

The measurement of physiological characteristics of esophageal voice has had a lot of attention during the last years. Part of this interest is due to the development of tracheo- esophageal valve prostheses [1]. Besides the advantages of these prostheses, a few disadvantages emerged too: the need to use one hand to close off the tracheostoma, the need for cleaning and exchanging the device, and the fact that relatively much effort is needed to phonate.

One more circomstance that leads to an interest from our side in this type of speech is the fact that in the Groningen ENT Clinic both injection- and button- esophageal voice are teached as a rule to the laryngectomees: this offers the opportunity to compare both types of speech on a physiological as well as on an evaluative level.

In the experiment reported on here this effort is assessed by measuring simultaneously the intra-tracheal, sub- and supra- pseudoglottic pressure, the trans- pseudoglottic flow, and the resulting sound pressure level. Furthermore, attention is payed to the pressure loss caused by the button, with the simultaneously measured air flow rate. The efficiency of voice production was measured, but it will not be reported on here. In fact, due to the relatively high intra- tracheal and sub- pseudoglottic pressures we encountered, these pressures will say as much about the effort of phonation as the efficiency.

The same speakers were asked to read a number of standard sentences. This speech material was judged by a group of 85 listeners, both naive judges and speech therapy students. The last group happened to consist of 96% female judges. Correlational computations have been made to relate these judgements to the physiological data of the same patients.

2) PHYSIOLOGICAL MEASUREMENTS.

In total, 1357 measurements were done in the phonations of 31 esophageal speakers. Not all variables were measured in every measuring point: during injection- esophageal phonations, we did not register the intra- tracheal pressure. The flow was measured in only 496 of the 1357 cases; this was done because the sound pressure level is influenced by the flow mask. The supra- pseudoglottic pressure data have not been processed so far.

The intra- tracheal pressure was measured with an open catheter, held by the patient himself in the trachea, under the thumb closing off the tracheostoma. The sub- and supra- pseudoglottic pressures were measured by means of micro pressure transducers, mounted on a catheter which has a diameter of 1.65 mm in the 6 cm of it between the two sensors. It was introduced through the nose into the esophagus, about 40 centimeters, and then gently pulled back again during phonation. By monitoring

the signal on a scope, evidence could be attained as to the position of the proximal sensor. When this sensor stops showing up pressure offset during phonation, it means that it is situated in the supra-pseudoglottic pharynx. Minor adjustments are sometimes needed in order to be sure that the distal sensor is situated in the air-filled sub- pseudoglottic room. The simultaneous registration of both sub- and supra- pseudoglottic pressure with high frequency sensors will enable us to investigate the acoustic phenomena occurring just below and above the pseudoglottis.

The spread in the data is quite high (see *Table 1*), especially in the sub- pseudoglottic pressure and the flow. The sound pressure level, on the other hand, has a rather small standard deviation, due to the generally small dynamic potential of these speakers.

When we consider the mean sub- pseudoglottic pressure, flow and SPL of our speakers it becomes clear which variables are able to differentiate between the two groups of injection- and button-esophageal speakers. The sub- pseudoglottic pressure seems to do that quite well. The means differ by 1.4 kPa. Four of the five patients where measurements were done during both types of phonation, showed a higher sub- pseudoglottic pressure (see *Table 2*). On the right hand side, you see the P-values from a comparison of the means with a t-test.

Table 1: Mean values and standard deviations of the physiological parameters; comparison of the button group and the injection group.

MEAN VALUES OF SUB-PSEUDOGLOTTIC PRESSURE, TRANS-PSEUDOGLOTTIC FLOW AND SOUND PRESSURE LEVEL.			
physiological variable	**all speakers**	**button group**	**injection group**
Psub (kPa)	3.3	4.1	2.6
(S.D.)	(1.6)	(2 2)	(2.2)
Flow (ml/s)	108	131	82
(S.D.)	(93)	(112)	(62)
SPL (dB(A))	66.6	66.2	67.0
(S.D.)	(9.5)	(9.4)	(9.7)

Table 2: Mean sub- pseudoglottic pressure values of 5 speakers, with the p-level of a t-test on difference of the means (between brackets).

MEAN SUB-PSEUDOGLOTTIC PRESSURE VALUES + T-TEST; 5 SPEAKERS.			
speaker	**button**	**injection**	**p<**
spr 4	2.8	< 3.5	.041
spr 9	3.9	> 3.1	.061
spr 11	5.3	> 2.2	.001
spr 24	5.4	> 3.1	.001
spr 34	2.6	> 2.4	.215

These differences ask for a physiological explanation. The question is: what will cause one and the same speaker to sustain two different pressures in order to vibrate one and the same sound source. In the first place the pressure build-up possibilities of the respiratory mechanism are responsible, allthough the pressure in the lungs and that in the sub- pseudoglottic room are not directly 1 to 1 related because the prosthesis is situated in between, and because the sub- pseudoglottic space is lying outside the thorax, hardly affected by the intra-thoracal pressure.

The flow values too differentiate between both types of speech. The mean registered value was 108 ml/s. Again, a significant difference was found between the two groups. Allthough we know from Schutte's data [2] that, for laryngeal voices, mean flow values are not very useful predictors of voice performance, the found differences might possibly be related to another voice variable: voice quality.

SPL did not discriminate between both voice types: both reached about 67 dB at 30 cm.

We measured a rather high in vivo trans-button pressure. These measurements were done without a selection based on the age of the devices. The mean age was about 11 weeks, so more than two and a half months; also, at the time we made our registrations (end 1985), no patients had anti-fungus medication [3]. The high trans-button pressures, with respect to the in vitro values we measured, must be attributed to the deterioration of the devices by fungal growth: it makes the material stiff, resulting in a higher flow resistance [3]. Consequently, research is going on at this moment to reduce the flow opposition of the prostheses.

3) PERCEPTUAL EVALUATION

Speech material of the same speakers was subjected to a perceptual evaluation by 85 listeners.. It was done by scoring one minute speech of each speaker on 13 semantic 7-points scales. The one minute speech was assembled out of the recorded material of read sentences. All pauses, coughs etc. were carefully cut out. We thought this useful to get right and reliable judgements, without the judges being distracted by all kinds of additional noises. In fact, it was not the noises that we wanted to evaluate, but the voices. The scales were 7-point scales. Of course we included in this experiment those 5 patients who were able to produce both types of esophageal voice.

Table 2: The set of 7-points semantic scales as used in the perceptual evaluation experiment (with English translation). The scales are adapted from Fagel et al., 1982.

SEMANTIC SCALES FOR VOICE EVALUATION

1)	zwak	— krachtig	(weak — powerful)
2)	onvast	— vast	(slack — firm)
3)	niet hees	— hees	(not husky — husky)
4)	eentonig	— melodieus	(monotonous — melodious)
5)	schel	— diep	(shrill — deep)
6)	traag	— vlot	(dragging — brisk)
7)	hortend	— vloeiend	(jerking — smooth-flowing)
8)	dof	— helder	(dull — clear)
9)	uitdrukkingsloos	— expressief	(expressionless — expressive)
10)	slecht verstaanbaar	— goed verstaanbaar	(not intelligible — intelligible)
11)	langzaam	— snel	(slow — quick)
12)	lelijk	— mooi	(ugly — beautiful)
13)	laag	— hoog	(low — high)

We performed factor analysis on the scores, and three main factors turned out to be important. As could be expected, the scales 6, 7 and 11 formed one

Tables 3 a, b: Multiple correlations computed with the speaker mean values on the 13 scales and the speaker mean values of three (or five in the case of button- esophageal speakers) parameters from the physiological measurements. *Flow* is the mean flow, *SPL* the mean sound pressure level, *Psub* the mean sub- pseudoglottic pressure, *Ptra* the mean intra-tracheal pressure, and *dPbu* the mean pressure difference over the buttons. The figures in italics point at statistically significant correlations ($p < .10$) as found during the multiple correlation steps. As soon as one or more of the three physiological parameters are in the equation, the other figures in the row point to correlations of the rest-variance. In rows without italic figures, Pearson product moment correlations are printed. The figures *in italics* of the last column (between brackets) are the squared multiple correlation values (with $p < .05$). Multiplied by 100 they give the percentage of the variation on that scale that can be predicted from the physiological measurement values.

Table 3 a: All speakers.

MULTIPLE CORRELATIONS *(Italics: p < .10)* COMPUTED WITH MEAN VALUES; ALL SPEAKERS.

scales	Flow	SPL	Psub	Multiple r (and r^2) *(Italics: p < .05)*
TEMPO factor:				
scale 6	*.46*	*.32*	-.01	*.55 (.30)*
scale 7	*.54*	.08	.11	*.55 (.30)*
scale 11	*.46*	*.33*	-.03	*.54 (.30)*
VOICE QUALITY factor:				
scale 1	*.50*	.14	-.03	*.52 (.27)*
scale 2	*.58*	.28	-.08	*.63 (.40)*
scale 3	-.11	-.02	.15	.19 (.03)
scale 4	*.66*	*.35*	-.13	*.71 (.51)*
scale 8	*.56*	*.47*	-.30	*.71 (.51)*
scale 9	*.66*	*.36*	-.16	*.73 (.53)*
scale 10	*.59*	.29	-.24	*.67 (.45)*
scale 12	*.62*	*.32*	-.29	*.70 (.49)*
PITCH factor:				
scale 5	-.22	*.58*	-.09	*.61 (.38)*
scale 13	*.43*	*.50*	-.11	*.63 (.39)*

Table 3 b: Button-esophageal speakers.

MULTIPLE CORRELATIONS *(Italics: p < .10)* COMPUTED WITH MEAN VALUES; BUTTON-ESOPHAGEAL SPEAKERS.

scales	Flow	SPL	Psub	Ptra	dPbu	Multiple r (and r^2) *(Italics: (p < .05)*
TEMPO factor:						
scale 6	.19	*.43*	*.59*	*-.46*	*.47*	*.83 (.69)*
scale 7	.10	.38	.33	*-.53*	*.62*	*.90 (.81)*
scale 11	.12	*.47*	*.45*	*-.49*	*.44*	*.79 (.62)*
VOICE QUALITY factor:						
scale 1	*.55*	*.49*	-.22	.05	.30	*.61 (.37)*
scale 2	.33	*.48*	.38	*.45*	*.61*	*.90 (.80)*
scale 3	-.17	-.37	-.35	-.03	-.13	.53 (.28)
scale 4	*.64*	.40	-.20	.00	.23	*.82 (.67)*
scale 8	*.44*	*.50*	*.54*	.17	.35	*.85 (.73)*
scale 9	*.65*	.37	-.25	-.00	.26	*.82 (.68)*
scale 10	*.57*	.41	-.47	.30	.37	*.86 (.74)*
scale 12	*.57*	*.52*	*-.49*	.25	.46	*.91 (.83)*
PITCH factor:						
scale 5	-.17	-.19	.09	.04	-.03	.32 (.11)
scale 13	.39	.25	-.32	-.06	.19	.56 (.31)

factor which one could call *tempo.* Another large factor can be considered to represent the *voice appreciation;* it comprises the scales 1, 2, 3, 4, 8, 9, 10 and 12. A third factor emerged as a *pitch factor:* scales 5 and 13.

This last factor showed a remarkable thing. As you can see, the scales have been arranged in such a way that the more negative side is on the left and the more positive side on the right. Now in the case of scale 5 it appeared that left and right should be changed. Normally, especially with male voices, the term "deep" would be more positive than "shrill". In the case of our set of esophageal speakers the reverse was the case. When the scale is turned around it correlates well with scale 13: "low - high", indicating that in fact "shrill" was considered a more positive attribute of the usually low pitched esophageal voices than "deep".

4) CORRELATION

Now the most interesting thing of such type of research is of course to correlate the physiological data on the one and the evaluative data on the other hand.

Considering all speakers *(Table 3 a)*, SPL and the "pitch" factor are relatively well correlated. The physiological relationship between these two variables in esophageal speech was established long ago. The mean flow values seem to be correlated with all factors, and especially with scales scoring on voice quality (e.g. scale 4 "melodious", scale 9 "expressive"). The same picture emerges from the correlations of the injection group. In the figures of *Table 3 b* the tendencies are somewhat different. The intra- tracheal and trans- button pressures correlate well with the "tempo" factor, and the "pitch" factor shows no important correlations at all.

One of the most striking things is that in all tables the mean flow shows relatively high correlations with the "quality" and "tempo" scales, and that the sub- pseudoglottic pressure hardly shows any correlation with the scales. We have no explanation as yet for this phenomenon, as one might expect this sub- pseudoglottic pressure to be an important determinant of voice quality; in any case, it is "closer to the voice" than e.g. the intra-tracheal pressure (see *Table 3 b*), but the latter has more to do with quality in the *Tables 3 a* and *b*.

REFERENCES

[1] H.F. Nijdam, A.A. Annyas, H.K. Schutte, H. Leever, «A new voice prosthesis for voice rehabilitation after laryngectomy», Clin. Otorhinolaryngol. 1982, 237: 27-33.

[2] H.K. Schutte, «The efficiency of voice production», Dissertation, Groningen, 1980.

[3] H.F. Mahieu, H.K.F. van Saene, H.J. Rosingh, H.K. Schutte, «Candida vegetations on silicone voice prostheses», Arch. Otolaryngol. Head Neck Surg. 1986, 112: 321-325.