THE EFFECT OF RADIOTHERAPY ON VARIOUS ACOUSTICAL, CLINICAL AND PERCEPTUAL PITCH MEASURES

Irma M. Verdonck-de Leeuw & Florien J. Koopmans-van Beinum
Institute of Phonetic Sciences/IFOTT, Amsterdam, The Netherlands

ABSTRACT

Speech samples of patients with early glottic carcinoma and of control speakers, are analysed for acoustical pitch and EGG, as well as by means of perceptual pitch evaluation by trained and untrained raters. The results of pitch ratings by trained listeners, EGG, and acoustical pitch correlate strongly and show the tendency that voices before radiotherapy have a higher pitch than 6 months and 2 years after radiation. Voices longer than 3.5 years after radiation tend to become higher again.

INTRODUCTION

Within the scope of a co-operative study between the Netherlands Cancer Institute (Antoni van Leeuwenhoek Hospital), the Academic Hospital of the Free University of Amsterdam, and the Institute of Phonetic Sciences of the University of Amsterdam, research is carried out on the effect of radiotherapy on voice quality. The aim of this study is to obtain parameters that can describe voice quality of patients with early glottic cancer before and after radiotherapy and of normal speakers. Voice quality can be described by several perceptual, clinical, as well as acoustical methods. In this presentation we will focus on various pitch measures. Data on acoustical pitch and EGG pitch are taken into account as 'objective' pitch measures; the results will be compared with perceptual evaluations of trained and untrained raters. The trained raters are used to provide an analytic description of voice quality. The role of the untrained raters is to find out how 'ordinary' people evaluate voice quality.

In a later stage of the study the results will be compared with other perceptual parameters of voice quality (evaluations on semantical scales, such as breathiness, harshness, creakiness), with other clinical methods (phonetogram, phonation quotient, and evaluation of stroboscopic recordings of vocal fold vibration), and with other acoustical analyses (LTAS, SNR, perturbation).

Before radiotherapy, the actual tumour can cause changes of the vocal folds such as mass change, stiffness change, and asymmetry. Little is known about voice quality after radiotherapy. Some studies report voice improvement to a normal or near-normal level, 6-12 months after radiotherapy, for about 70% of the patients [1,2,3]. Other studies report abnormal post-radiation voices [4,5]. Pitch may be one of the parameters that can be influenced by the presence of a tumour or by the effect of radiotherapy on the vocal fold tissue, such as late oedema, necrosis etc. Furthermore, pitch measurements are important cues for other acoustical and perceptual measurements, such as spectral noise [6,7,8], breathiness, and tension [8,9].

METHOD

Speakers/recordings

Patients with early glottic cancer (T1N0M0) are treated with radiotherapy (60 Gy in 30 fractions, or 66 Gy in 33 fractions). Voice samples of the same 10 patients have been recorded before radiation, as well as 6 months, and 2 years after radiation. Recordings are also made of 3 other groups of 10 patients each, before radiation, 6 months, and 2 years after radiation, and of 20 patients longer than 3.5 years after radiation. Finally, recordings are made of 20 control speakers without any known vocal defects (figure 1). The matching between patients and control speakers took into account sex (all male), age (47-81), as well as smoking habits. The speakers read aloud a text for about 5 minutes. All the material was recorded using a Casio DAT-recorder and a Philips N8214 microphone. Fragments (ca 30 sec.) of all texts were digitised by means of the Sound editor of an Iris Indigo R4000, sample frequency 48 kHz, 16 bit resolution. These samples were copied to cassette tapes in two random speaker orders.

	PRE	P1	P2	P3	
Long.	10 ->	10 ->	10		
Mix.	10	10	10	20	
Control					20

Figure 1. Illustration of the various speaker groups. Longitudinal group before radiation (PRE), 6 months after (P1), and 2 years after (P2). Mixed groups PRE, P1, P2 and P3 (longer than 3.5 years after radiation). Control group.

Raters/rating procedure

The untrained raters in this experiment are 20 university students (6 male, 12 female), without any experience for this listening task. They were paid for their participation. The raters received written instructions. First they heard examples of 10 different voices in order to get a reference frame. After the examples 110 fragments of read aloud text were presented (10 training fragments and 100 fragments of speakers as indicated in figure 1). The raters judged voice quality on 14 voice quality scales. The tapes were presented binaurally via a cassette recorder and headphones. The raters listened to the tapes in a quiet room, individually. On the average, the whole rating procedure (instructions + rating) took about 1 1/2 hours.

The trained raters are 3 female (socio-) phonetic researchers; 2 had followed a training course on the Voice Profile by John Laver. They rated the voices on 8 scales independently from each other.

The semantic scales consist of various voice quality scales that have been used in previous experiments [9]. In this paper we limit ourselves to the 'pitch' scales *low-high* and *deep-shrill* (7-points) used by the untrained raters, and *low-high* and *sonor-shrill* (13-points) used by the trained listeners.

Acoustical Pitch

Average pitch values of the read aloud text are determined by means of the program "Praat" [10]. The acoustic pitch period of a sound is determined by the position of the maximum of the autocorrelation function of the sound. This procedure is extensively described by Boersma [10]. In order to select only voiced candidates for pitch detection the Voicing Threshold is set to 0.5 and the Silence Threshold to 0.05. All other parameters are kept default.

Clinical fundamental frequency

By means of an electroglottograph (Stopler Teltec GFA06) the average fundamental frequency is measured for the read aloud text. The speakers read aloud the same text for about 5 minutes while 1000 voiced samples were analysed and averaged.

RESULTS

The reliability coefficient Ru is used as a measure of the reliability of the means of the ratings by a panel of raters (between 0 and 1). Ru= (MSspeakers - MSraters) / MSspeakers. The results for the 3 trained raters are Ru = .80 for *low-high* and Ru = .63 for *shrill-sonor*; for the 20 untrained raters Ru = .93 for *low-high* and Ru = .94 for *deep-shrill*. The differences between trained and untrained raters lies in the low MSspeakers (=true variance) by the trained raters.

Agreement is determined by Kendall's W (between 0 and 1). For the trained raters W = 0.68 for *low-high* and W= 0.64 for *sonor-shrill*; for the untrained raters W = 0.35 for *high-low* and W= 0.30 for *deep-shrill*.

Means of the ratings per speaker per scale for the trained raters, and for the untrained raters are put into a datafile, together with the acoustical pitch data, as well as the EGG data (figure 2). Pearson correlations for the various means are given in table 1.

	2	3	4	5	6
1	.89	.46	.11	.45	.42
2		.55	.08	.57	.50
3			.24	.72	.69
4				.30	.16
5					.75

Table 1. Pearson correlations for the 6 pitch measures. 1.'low-high' untrained, 2. 'deep-shrill' untrained, 3. 'low-high' trained, 4. 'sonor-shrill' trained, 5. acoustical pitch, 6. EGG

The results of the individual data given in figure 2 show that the perceptual evaluations by trained and untrained raters do not differentiate between the speaker groups. Over all speakers, the trained raters range from 7.0 to 7.8 for *low-high* and from 6.5 to 7.4 for *sonor-shrill*. The untrained raters range from 2.7 to 3.3 for

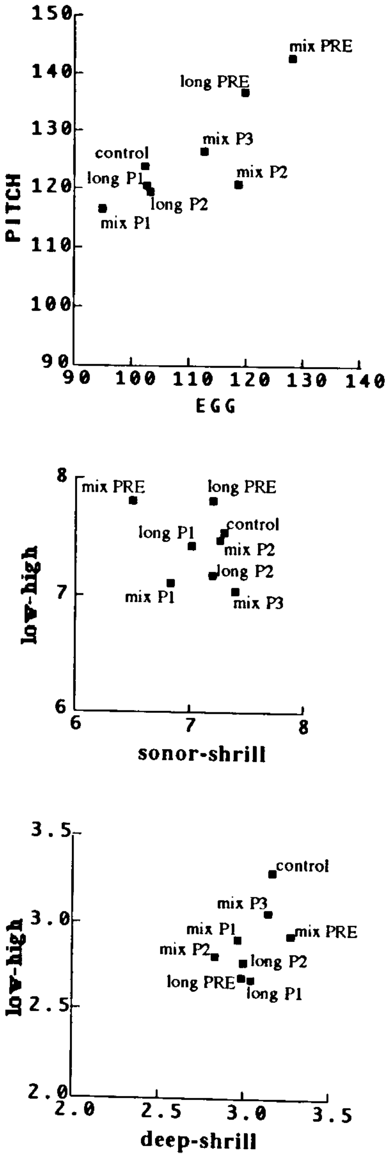

Figure 2. Means of acoustical pitch by EGG (top), and 'low-high' by 'sonor-shrill by trained raters (middle), and 'low-high' by 'deep-shrill' by untrained raters (bottom) for Longitudinal groups before radiation (PRE), 6 months after (P1), and 2 years after (P2), and Mixed groups PRE, P1, P2, and P3 (longer than 3.5 years after radiation), and control group.

low-high and from 2.8 to 3.3 for *deep-shrill.* Notice the deviating scales in figure 2: 6-8 for the trained raters (13 points scale), and 2-3.5 for the untrained raters (7-points scale).

It is obvious that the raters didn't hear specific differences between the speaker groups. This was already indicated by the reliability and agreement results: the trained raters having a high agreement score but a low reliability coefficient, due to a low MSspeakers.

The results for the acoustical pitch and EGG data do show (though statistically not significant) differences between the speaker groups. To find out wether a combination of parameters will give better insight in the effect of radiation on pitch a factor analysis is carried out.

A Principal Component Analysis is used to decompose the correlation matrix into (varimax rotated) factors (PCA). For determining the number of factors, the criterion 'eigenvalue > 1' is applied. With this criterion the PCA produced 2 factors, together explaining 80% of the total variance. On the basis of the factor loadings (> .6) the 2 factors are mainly determined by acoustical pitch, EGG, and *low-high* by trained raters (factor 1), and *low-high* and *deep-shrill* by untrained raters (factor 2). Factor scores are presented here for the first factor as it explains most of the variance (49%). The scores give the position for each speaker on each factor (figure 3). Results show that a tendency can be seen (though statistically not significant) for 'pitch' as a combination of acoustical pitch, EGG and *low-high* evaluations by trained raters: patients before radiotherapy have a very high 'pitched' voice as compared to patients 6 months and 2 years after radiation. This counts for both patient groups (longitudinal and mixed). The voices of patients longer than 3.5 years after radiation tend to become higher again, whereas the control speakers have the lowest voices.

The tendency that patients before radiation have very high pitched voices may be due to mechanical effects of the tumour on the vocal folds. Another explanation may be an increased tension of the vocal folds by the patient in order to compensate for his voice loss. Also, little is known about the effect of microlaryngeal

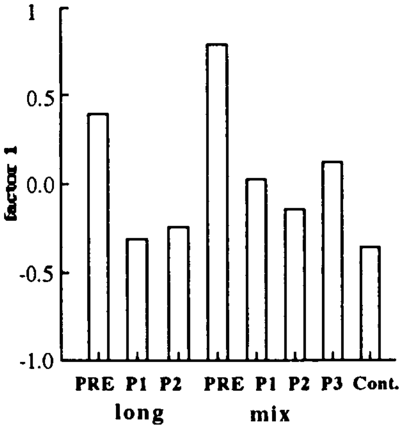

Figure 3. Means of Factor 1 for Longitudinal groups before radiation (PRE), 6 months after (P1), and 2 years after (P2), and Mixed groups PRE, P1, P2, and P3 (longer than 3.5 years after radiation), and control group.

surgery that most of the patients have undergone before radiation.

After radiotherapy the voices seem to be low pitched. Conflicting results have been found in previous research [4,9], though none of the results were significant. On the long term, the effect of radiotherapy seems to be that voices tend to become 'normal' again (2 years after radiation) but become high pitched later on. This may be due to the irradiation of the normal tissues that can result in late oedema, altering the vibratory cycle by mass and stiffness changes of the vocal folds.

Although the results in this experiment do not differentiate significantly between the speaker groups, the correlations between the acoustical pitch analysis, and the EGG data, and the pitch evaluations of the trained raters for the read aloud text are clear. The expectation that what one can hear should also be measurable, becomes true in this experiment, at least for the trained raters.

The evaluations by the untrained raters do not correlate strongly with the other analyses; still they do have an important role in this study. The purpose is to find out how 'ordinary' people (i.e. not voice reseachers/pathologists) evaluate voices of patients. Therefore in future research, the evaluations of the patients themselves and their partners are taken into account as well.

The expectation is that raters, trained as well untrained, can differentiate between speaker groups on other aspects than pitch evaluations. Obviously, pitch is not a parameter that represents strongly the effect of radiotherapy, since none of the various pitch measures discriminates clearly between the speaker groups.

REFERENCES

[1] Harrison, L.B, Solomon, B., Miller, S., Fass, D.E., Armstrong, J. & Sessions, R.B. (1990). "Prospective computer-assisted voice analysis for patients with early stage glottic cancer: a preliminary report of the functional result of laryngeal irradiation", *Int. J. Radiation Oncology Biol. Phys.* 19, 123-127.

[2] Miller, S., Harrison, L.B., Solomon, B. & Sessions, R.B. (1990). "Vocal changes in patients undergoing radiation therapy for glottic carcinoma", *Laryngoscope* 100, 603-606.

[3] Lehman, J.J., Bless, D.M. & Brandenburg, J.H. (1988). "An objective assessment of voice production after radiation therapy for stage 1 squamous cell carcinoma of the glottis", *Otolaryngol Head Neck Surg* 98, 121-129.

[4] Colton R. H., Sagerman, R.H., Chung, C.T., Yu, Y.W. & Reed, G.F. (1978). "Voice change after radiotherapy", *Radiology* 127, 821-824.

[5] Emanuel, F.W. & Smith, W.F. (1974). "Pitch effects on vowel roughness and spectral noise", *Journal of Phonetics* 2, 247-253.

[6] Wolfe, V., Cornell, R. & Palmer, C. (1991). "Acoustic correlates of pathologic voice types", *Journal of Speech and Hearing Research*, 34, 509-516.

[7] Verdonck-de Leeuw, I.M. & Boersma, P. (to appear). "The effect of radiotherapy measured by means of Harmonicity".

[8] Hammarberg, B., Fritzell, B., Gauffin, J., Sundberg, J. & Wendin, L. (1980). "Perceptual and acoustic correlates of abnormal voice qualities". *Acta Otolayrngologica* 90, 442-451.

[9] Leeuw de, I.M. (1990). "The relation between perceptual and clinical parameters of voice quality of patients with early glottic cancer before and after radiotherapy and of normal speakers", *Proceedings of the Institute of Phonetic Sciences Amsterdam* 14, 27-38.

[10] Boersma, P. (1993). "Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound", *Proceedings of the Institute of Phonetic Sciences Amsterdam* 17, 97-110.

[11] Hoyt, D.J., Lettinga, J.W., Leopold, K.A. & Fisher, A.R. (1992). "The effect of head and neck radiation therapy on voice quality", *Laryngoscope* 102, 477-480.