Acoustic Measurement of Voice Quality in Dysphonia after Traumatic Midbrain Damage

E. Hartmann and D.v.Cramon
Munich, Federal Republic of Germany

1. Introduction

The advantage of acoustic measurement of voice disorders in clinical investigation and treatment has often been emphasized.

Basic categories of pathological voice quality have been established by auditory judgement. However, often there is little agreement on these categories regarding their relation to the various vocal dysfunctions.

In a recent study (Cramon and Vogel, in press) the perceptual categories 'breathy,' 'rough' and 'tense' proved to be sufficient to describe the features of central dysphonia. This is substantiated by assignment of these categories to three basic laryngeal settings (Laver, 1980). Breathy voice quality is characterized by insufficient adduction of the vocal folds during phonation. The acoustic correlate is an increase in the spectral noise components. Rough voice quality is caused by irregularities in the vibration of the vocal folds, due to rigidity or asymmetric tension. This correlates acoustically with aperiodicity of the fundamental frequency. Tense voice quality is due to hyperadduction of the vocal folds and hypertension in the walls of the supralaryngeal cavities. The concomitant acoustic feature is an increase in the upper harmonic components in the spectrum.

Our main aim was to differentiate the above mentioned pathological voice qualities acoustically and to make available an index for the degree of severity of these pathologies.

2. Method

14 male and 10 female patients were examined. 18 of them had suffered from a severe closed head trauma with subsequent traumatic midbrain syndrome and the remaining 6 patients had suffered from a cerebro-vascular accident. Phonetic testing had revealed general symptoms of dysphonia in all patients, such as reduced range of pitch and loudness, increased pitch level and deviant voice quality. They were chosen from a larger sample so as to form three groups of equal quality. The presence of peripheral lesions could be excluded following phoniatric examination. 10 males and 7 females, age-matched, served as control group.

Speech material consisted of the cardinal vowels /u, o, a, e, i/ uttered twice by each subject. The recorded items were digitized and the fundamental periods were tracked using the autocorrelation method. Subsequently the voiced segments of each vowel were subjected to spectral analysis.

On the basis of these data 5 parameters were computed. The parameter 'Time lag of pre-exhalation' Tb was defined as the duration of the pre-exhalation noise. A noise energy threshold was used to determine automatically the beginning of turbulent airflow, and the first fundamental period signalled the end of the pre-exhalation segment. The parameter 'Fundamental period perturbation' PP measured the average percentual deviations of consecutive fundamental periods from a smoothed trend line. In the frequency domain, the parameters A1 and A5 measured the relative amount of spectral energy in the frequency bands from 1 to 5 KHz and from 5 to 10 KHz respectively. V5 was calculated as the spectral variance in the range from 5 to 10 KHz, after band-pass filtering of the spectral function. The described parameters were computed for each of the vowels and were subsequently averaged.

3. Results

The results for the parameter Tb indicated discrimination of the group of breathy voices from each of the other groups, which was highly significant. The breathy voices revealed Tb values between 40 and 120 msec. Only one of the control speakers showed a notable pre-exhalation of 10 msec duration, 3 of the tense voices reached this upper limit and little increase of Tb was observed for 5 subjects with rough voices.

The parameter PP differentiated significantly the group of rough voices from each of the other groups. Rough voices exhibited PP values between 1.6 and 2.6 %. These results are lower than those measured in peripheral lesions. Three tense voices with increased PP may have additional roughness, thus a slightly harsh voice. One tense voice demonstrated PP values even below the control range. It showed a vibration mode which may be characterized as 'mechanical.' Comparing Tb and PP values of some subjects, we may assume a hoarse voice type, if both parameters exhibit high values.

The scatter diagram of the spectral parameters A1 and A5 displayed the control group at the lower left corner with low A1 and A5. The values for breathy and tense voices tended to shift up and/or to the right with tense voices (except for one subject) showing a lower A5 than the breathy voices. Discriminant analysis separated breathy and tense voices significantly from each other as well as from the control group. The rough voices lay in the range of the normal voices in most of the cases.

The particular distribution of spectral energy for breathy voice quality reflects an overall increase of noise energy with a predominance in the 5 to 10 KHz range. Thus the parameter A5 seems to be an additional parameter for classifying breathy voice quality. A significant correlation of parameter Tb and A5 supported this assumption. The particular form of the spectral distribution of tense voice quality may be due to an increase in higher harmonic components with a predominance in the 1 to 5 KHz range. The tense voice with high A5 may be a compound voice type with additional breathiness.

The parameter V5 turned out to be an additional acoustic correlate of the perceptual category 'tense voice.' However, classification was only reliable for male tense voices. Tense voice quality seems to be produced not only by laryngeal but also by supralaryngeal settings. Overall tension with hard surfaced vocal tract walls may improve the resonance and narrow the formant bandwidths. This may cause an increased variance of the spectral amplitude.

In a follow-up study the voices of 13 patients with dysphonia after traumatic midbrain damage were measured at 2 different stages. (Stage I: 1-12 weeks after mutism; Stage II: more than 24 weeks after mutism). Additionally to the discussed parameters we calculated the mean fundamental frequency MF_0.

2 subgroups of patients were established. The first subgroup consisted of 5 patients who revealed spectral parameters in the normal range at stage I. The second subgroup consisted of 8 patients who initially showed abnormally high spectral parameter values.

At stage II the first group exhibited an increase of the spectral parameters A1 above the range of the control speakers/ whereas the parameter A5 remained almost constant in the normal range. This particular relation of A1 and A5 indicated evolving of tense voice quality.

The second group exhibited a decrease in A1 and A5, which had been particularly high at the initial stage. This indicated a decrease of breathy voice quality. Both groups showed decreasing values of the parameters MF_0 and Tb. These findings were in accordance with the auditory judgement: a decrease in initially high pitch and breathiness, and evolving of tense voice.

The group means of PP and V5 lay in the normal range at both stages. These parameters only indicated individual changes.

4. Conclusion

In summary we found objective acoustic measures for the auditory categories of pathological voice quality, which proved to be appropriate to separate the components of compound voice types. Moreover, different stages in the process of phonatory recovery could be described quantitatively.

References:

Cramon, D., von and Vogel, M. (in press). *Zentrale Stimmstoerungen*. Erlangen: Perimed-Verlag.

Laver, J. (1980). *The phonetic description of voice quality*. Cambridge: Cambridge University Press.