Objective measurement of the severity of dysphonia typically requires signal processing algorithms applied to acoustic recordings. Since Lieberman (1963) introduced the concept of perturbation analysis in the area of voice, the most dominant acoustic parameter in clinical practice is the classical jitter. However jitter measurements have some critical limitations. According to a widely accepted guideline, in sustained vowels of dysphonic voices, only perturbation measures less than about 5% (quasi-periodic voices) are reliable: this is related to period extraction methods.

This means that traditional acoustic analysis programs available for clinical use are not suited for quality assessment of strongly irregular voices, as substitution voices (voices not generated by two vocal folds, particularly after total/partial laryngectomy) or spasmodic dysphonias. The basic protocol for multidimensional voice assessment as recommended by the European Laryngological Society (Dejonckere et al., 2001) specifically mentions that it is not suitable for a few very special categories of voices, as substitution voices and spasmodic dysphonias. Nevertheless a valid quality evaluation is essential for substitution voices, as in laryngeal oncology there may be different therapeutical options comparable in survival rate for the same nature and stage of cancer. In such cases, functional outcomes (voice, respiration, swallowing) gain major significance. The strong irregularity that characterizes the substitution voices is the major problem for usual acoustic analysis.

This special session deals with a part with successful improvements of the traditional approach of the cycle-to-cycle variability. A breakthrough was made possible by the development of a synthesizer of ‘realistic’ pathologic voices, that cannot be recognized by expert listeners from true patient’s voices, and where the jitter ‘put in’ is exactly known. This allows to check as well the ability of pattern recognition of the human visual system as the validity of new algorithms for period detection, in different conditions of noise. The practical result is that the traditional threshold limit value of 5% for jitter measures may be transgressed under some conditions that will be discussed.

Furthermore, the question remains about the clinical value of perturbation measurements when analyzing running speech of patients with either substitution voices or spasmodic dysphonia. The same question is relevant for noise measurements. It still becomes clearer that the acoustic parameters that are in some way related to the selection of voiced/unvoiced parts of the signal are the most successful ones in discriminating either different types of substitution voices or therapeutic effect in spasmodic dysphonia.

Another problem is the presence of tremor in some pathological voices: this mainly concerns neurological voices, and particularly spasmodic dysphonia, a focal laryngeal dystonia. The estimation of tremor attributes in a speech signal involves the accurate extraction of the signal that modulates the time-varying fundamental frequency. A new significant attribute of tremor is introduced. It derives from the time-varying characteristic of the modulation level, namely the deviation of the modulation level. The mean modulation level and its deviation are combined in a quality indicator trying to classify speakers according to the prevalence of tremor in their voice. This innovative approach can be tested on sustained vowels uttered by patients who suffer from spasmodic dysphonia before and after medical treatment.

REFERENCES


Dejonckere PH, Giordano A, Schoentgen J, Fraj S, Bocchi L, Manfredi C “Validity of jitter measures in non...
