

ACOUSTIC ASSESSMENT OF NEUROGENIC VOICE DISORDERS IN A CLINICAL SETTING

Michael Merk⁺, Wolfram Ziegler^{*}, Bettina Brendel^{*}

⁺ Inst. for Communication Engineering, Fed. Armed Forces University, München, Germany

- EKN Clinical Neuropsychology Research Group, Dachauer Str. 164, D-80992 München, Germany. E-mail wolfram.ziegler@extern.lrz-muenchen.de

ABSTRACT

A new tool for the computerized clinical assessment of neurogenic voice disorders is presented. It is part of a more comprehensive PC-based system for the diagnosis of speech impairments in patients with neurologic disorders, including modules for the assessment of speech rate, fluency, and articulation, in addition to the voice module. The design of this instrument is optimized with regard to stability and precision rather than computational speed. The voice analysis parameters used to-date are fundamental frequency and its variation over time, intensity and its variation over time, spectral tilt, and the cepstrum. The system controls all diagnostic steps, from the presentation of stimulus materials to the evaluation of results. Although the procedure is highly automatized, it allows for interactive control of recording quality, handling of algorithm instabilities, and correction of implausible results.

INTRODUCTION

Virtually all neurologic diseases interfering with general motor control can also involve laryngeal motor functions and, as a consequence, lead to voice disorders.

Usually, neurogenic dysphonia is part of a more general dysarthric disorder which also includes motor problems of the respiratory and the supralaryngeal muscles during speaking. The voice problems encountered in these patients may either reflect direct involvement of the laryngeal motor system or a compensatory reaction to respiratory or articulatory dysfunctions [1].

Clinical assessment of neurogenic voice disorders is in the first instance based on perceptual evaluations of the patient's voice [2]. This method, however, is inherently subjective and, in most cases, not sufficiently reliable for the purpose of assessing spontaneous or therapeutically induced change. More objective methods of evaluating vocal performance are based on speech signal analysis techniques.

Although there is abundant literature on acoustic measures of vocal pitch, voice quality, and voice stability, and although several measurement systems are commercially available, many clinicians are still hesitant in actually using these

measures in their routine work. One of the reasons for their reluctance may be the fact that most systems are too unconstrained and fail to provide a stable framework, from suitable assessment protocols over recording tools and interactive analysis methods to comprehensive tabular and graphical presentations of the relevant data. The system presented here tries to provide a user-friendly assessment tool for use in the diagnosis of neurogenic speech disorders. The assessment protocol encompasses sentence repetition, sustained vowel production, word repetition, and articulatory diadochokinesis. Acoustic variables are focussed on speech rate, fluency, consonant articulation, vowel articulation and *voice*. Several design principles were adopted to increase the acceptance of the system in everyday clinical applications: Signal analysis methods were optimized on precision rather than computational speed. No manual segmentation is required. Results can be checked for their plausibility and can be accepted, corrected, or rejected on different stages of the analysis process. Individual tasks can easily be repeated in cases of questionable data. The results are presented graphically or in tabular form.

The system is based on a PC with a commercial sound card and uses MATLAB as programming environment. The design concept allows for rapid prototyping, fast code generation and relieved algorithm fitting after clinical tests [3].

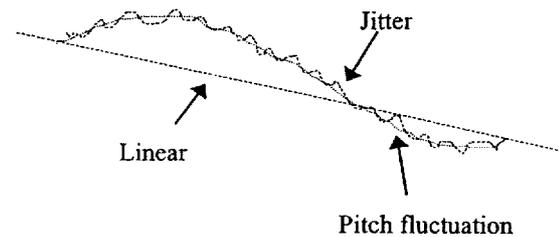
VOICE ANALYSIS

This module is based on sustained vowel production tasks based on the German vowels /i/ and /a/. The patient is required to produce each vowel several times over a few seconds at a comfortable pitch and intensity level. These speech samples are also evaluated in the vowel articulation module of the system.

Acoustic evaluation of the patient's utterances is performed off-line. Mid-portions of the sustained vowels are segmented automatically on the basis of psychoacoustic loudness contours. Fundamental periods are then detected by an adaptive tracking algorithm which uses a cepstrum-based estimate of the mean F0 as its starting value. Manual specification of alternative starting values can be used to obtain valid trajectories even in cases of poor voice quality or other abnormalities. In a subsequent step, pitch and intensity trajectories over time are analysed with regard to their variability on three different time scales:

- (1) Fast oscillations of the order of single pitch cycles, i.e. jitter and shimmer. Variability of this type can be caused by aberrant vocal fold tension, leading to irregularities in the duration and amplitude of vocal fold oscillations.
- (2) Slower variations, of the order of several pitch cycles, referred to as *pitch* or *intensity fluctuations*. These irregularities may result from intermittent alterations in vocal fold tension, such as in dystonic or ataxic disorders [1].
- (3) Linear trends of F0 or intensity, e.g. caused by vocal fatigue.

Figure 1 illustrates how these parameters are extracted stepwise from the pitch or intensity contours.



As a measure of *breathiness*, a frequency

Figure 1: Extraction of different measures of F0-variation

domain measure of signal periodicity (Cepstral peak prominence smoothed, CPPS) is used [4]. CPPS contours over time are computed automatically over sustained vowel sections. They can be used to visualize intermittent alterations in voice quality, such as in spasmodic dysphonia (see figure 2). Corresponding high-resolution-power-spectra may provide additional information about signal periodicity and spectral tilt.

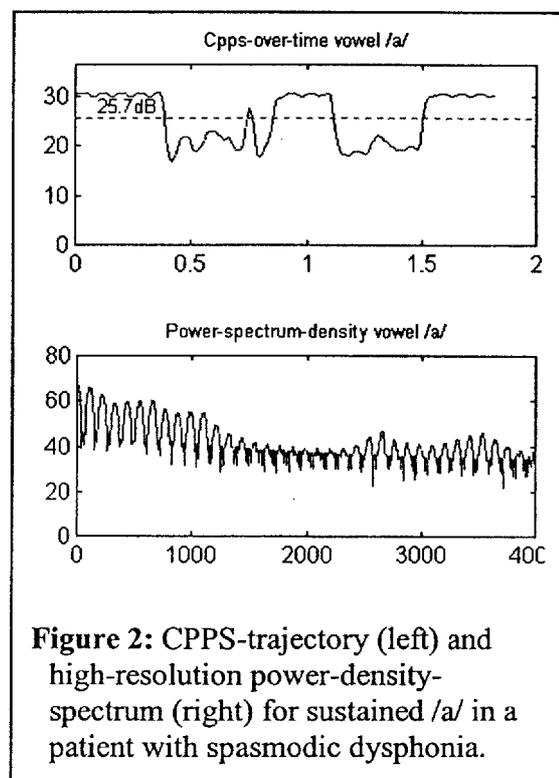


Figure 2: CPPS-trajectory (left) and high-resolution power-density-spectrum (right) for sustained /a/ in a patient with spasmodic dysphonia.

APPLICATIONS

The effectiveness of the described acoustic measures in predicting perceptual roughness and breathiness ratings was tested on a set of vocoder voices. Rosenberg's glottal source function was used with user-defined random jitter values (simulation of roughness) and/or additive white noise with predefined signal-to-noise-ratios (simulation of breathiness). Significant correlations were obtained between model parameters and the diagnostic measures proposed here. The system is currently being applied in a clinical setting. First results for a group of patients with neurogenic speech disorders confirmed the clinical applicability of the system and the sensitivity of its parameters in detecting neurogenic voice problems.

REFERENCES

- [1] Ziegler, W. and Hoole, P. (1999). Neurologic disease. In R.D. Kent & M. Ball (Eds.), *Handbook of Voice Quality Measurement*. San Diego: Singular Publishing Group.
- [2] Darley, F.L., Aronson, A.E. and Brown, J.R. (1975). *Motor Speech Disorders*. Philadelphia: W.B. Saunders.
- [3] Merk, M. and Ziegler, W. (1999). MoDiaS – a PC-based system for routine acoustic analysis of neurogenic speech disorders. In B. Maassen & P. Groenen (Eds.), *Pathologies of Speech and Language*. London: Whurr Publishers.
- [4] Hillenbrand, J. and Houde, R.A. (1996). Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *Journal of Speech and Hearing Research*, **39**, 311-321.