ASSESSMENT OF DYSFLUENCY IN STUTTERED SPEECH

Bergl P.[1], Lustyk T.[1], Čmejla R. [1]
Černý L.[2], Hrbková M.[2]

[1] Faculty of Electrical Engineering, Czech Technical University in Prague,
[2] Department of Phoniatrics, 1st Faculty of Medicine Charles University and General Faculty
Hospital in Prague

Abstract
The project aims to find methods that would be based on the analysis of audio recordings stutterers and were able to objectively and automatically determine the degree of speech fluency disorders.

1 Database

All experiments were performed on speech signals from 121 speakers, recorded in the past few years at Department of Phoniatrics, 1st Faculty of Medicine Charles University in Prague. The youngest speaker was 7 years and 5 months, the oldest 49 years and 5 months. Women and girls represent 23%. Subjective assessment of stuttering and dysfluency has been conducted in 2008 by two clinicians at Phoniatrics clinic in Prague.

2 Methodology

The analysis of annotations from stutterer's spontaneous speech has been done firstly. Also the annotations were processed by recognition HTK system. Dysfluency occurence was also analyzed in the groups of parameters in time domain the ratio of the total length of silence and speech, the number of segments of silence/speech, the number of ascending edges of energy envelope, the periodicity of speech signal energy, parameter voice/unvoice speech and other original parameters) and in the spectral domain (our original approach based on the detection of significant spectral changes using Bayesian detector, GLR distance and HTK recognizer). The effect of ambient noise in the evaluation was also analyzed.
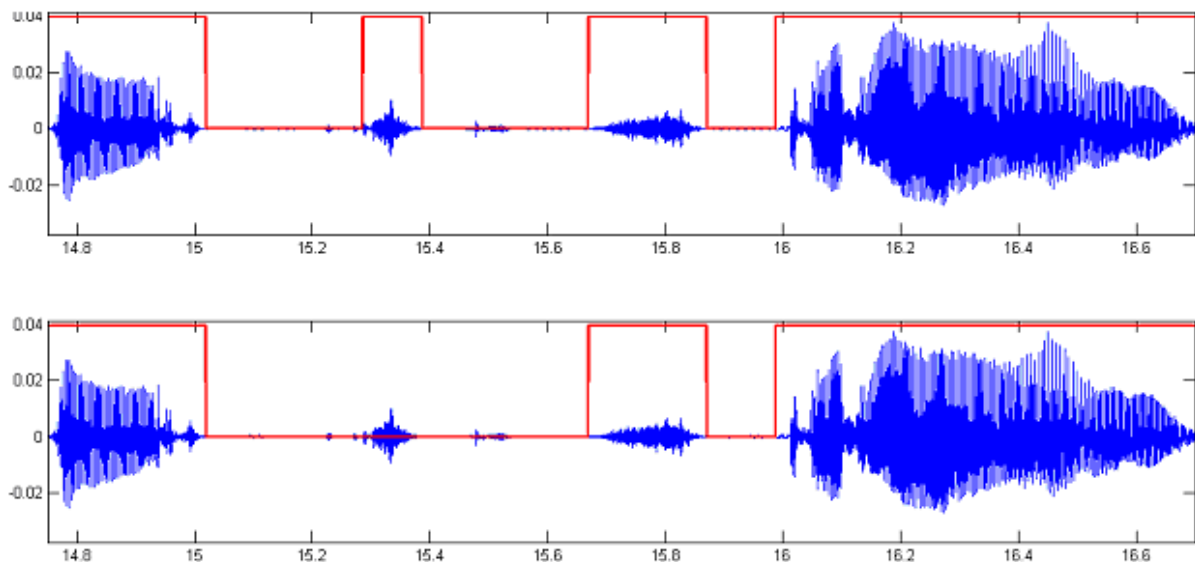

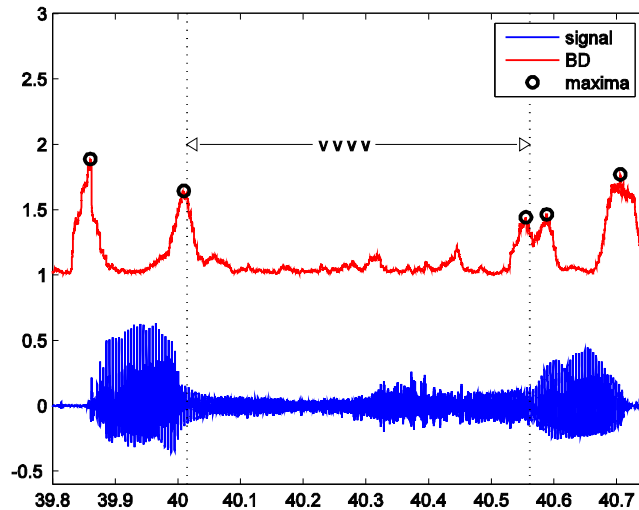Figure 1: Example of signal processing in time domain.

Figure 2: Example of signal processing in frequency domain.

All trials were analyzed in detail in terms of deviations from the evaluation of the two phoniatric experts. It was found that the number of repetitions and prolongation together strongest correlates with the standard deviation of the moving average of spectral changes. The number of disconnected words strongest correlates with spectral changes computed for the 80% quantile. Detailed results for all parameters are given in Table 1 and Table 2.

time domain	
parameter	correlation coefficient
the average length of silence	0.793
the periodicity of speech signal energy	0.767
the ratio of the total length of silence	0.730
the parameter voice/unvoice speech	0.711
the occupancy levels of the speech signal	0.655
the number of segments of silence and speech	0.616
the number of ascending edges of energy envelope	0.592

Table 1: Correlation coefficient of parameters in time domain and the assessment of clinicians.

frequency domain	
parameter	correlation coefficient
the average number of maxima - HTK	-0.795
the average number of maxima - BACD	-0.782
the average number of maxima - GLR	-0.778
the standard deviation distance from MA - BACD	0.768
the standard deviation distance from MA - HTK	0.767
the standard deviation of distance maxima - HTK	0.762
the standard deviation distance from MA - GLR	0.761
the standard deviation of distance maxima - BACD	0.757
the standard deviation of distance maxima - GLR	0.745
the 80% quartile distance of peaks - BACD	0.737
the 80% quartile distance of peaks - GLR	0.733

Table 2: Correlation coefficient of parameters in frequency domain and the assessment of clinicians.

The proposed system of objective evaluation, combining a large number of measurements. It is advantageous to combine parameters that reflected various kind of dysfluency. Better results are achieved by combination of suitable parameters than using individual parameters. Combination method is illustrated in Figure 3.
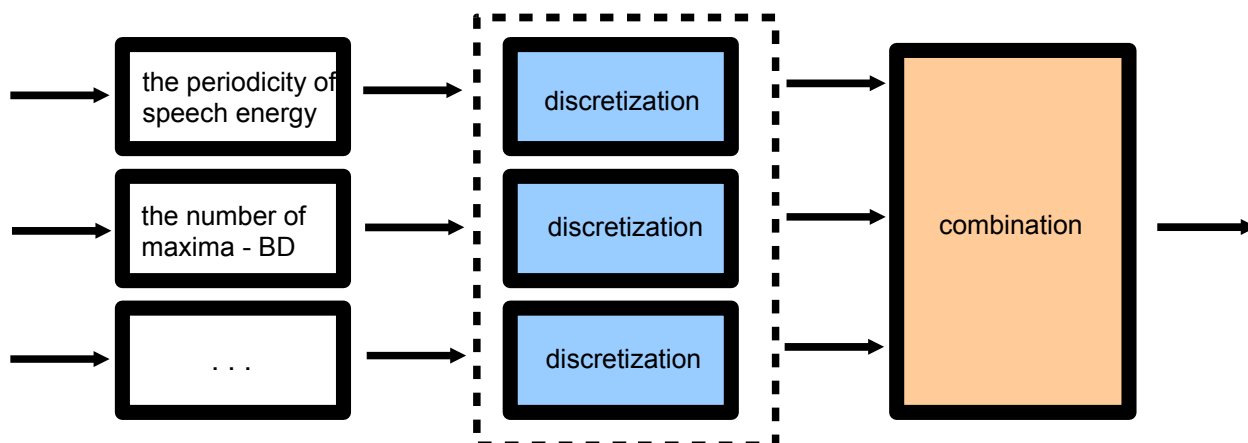

Figure 3: Combination method for a large number of parameters.

3   Conclusion

All trials were analyzed in detail in terms of deviations from the evaluation of the two phoniatric experts. It was found that the number of repetitions and prolongation together strongest correlates with the standard deviation of the moving average of spectral changes. The number of disconnected words strongest correlates with spectral changes computed for the 80% quantile.

The proposed system of objective evaluation, combining a large number of measurements, correctly estimated the degree of dysfluency for 63% of individuals with an error in the estimation of more than one class occurs only in 1.7% of speakers.

Acknowledgment

References

[1] Bergl P., *Objektivizace poruch plynulosti řeči*. Dissertation thesis, Faculty of Electrical Engineering, CTU in Prague, 2010.

[2] Bergl P., Čmejla R., Hrbková M., Černý L.: *System for automatic evaluation of stuttering*. Automatizace 2010, roč. 53, č. 1-2, s. 49-52. ISSN 0005-125X. (in Czech).

[3] Lustyk T., *Analýza neplynulé řeči*. Diploma thesis, Faculty of Electrical Engineering, CTU in Prague, 2010.

Bergl P.   berglpet@fel.cvut.cz

Lustyk T.   lustytom@fel.cvut.cz

Čmejla R.   cmejla@fel.cvut.cz