VOICE CEPSTRAL ANALYSIS IN ADDUCTOR-TYPE SPASMODIC DYSPHONIA

G. Costantini¹, G. Saggio¹, D. Casali¹, L. Marsili², A. Suppa², A. Berardelli², G. Ruoppolo³

¹Department of Electronic Engineering, University of Rome Tor Vergata
Via del Politecnico 1, 00133 - Rome, Italy
²Department of Neurology and Psychiatry and, Sapienza University of Rome
³Department of Sensorial Organs, Otorhinolaryngology Section, Sapienza University of Rome

Adductor-type spasmodic dysphonia (ASD) is a task-specific focal dystonia manifesting with involuntary laryngeal muscle spasms leading to intermittent strained/strangled voice [1] (Fig. 1). ASD is often poorly recognized by clinicians not familiar with the disorder, because of the lack of diagnostic criteria and of validated severity scales. Following our recently published observations [2], we performed voice analysis in ASD patients by using cepstral analysis. Cepstral analysis is based on Fourier transform of the logarithm power spectrum of an acoustic signal and reflects the dominant rahmonic in the voice sample [3].

Fig. 1 - Upper panel: laryngoscopic exam of vocal folds in ASD during phonation (right) and during rest (left).
Lower panel: narrow band spectrogram: example of sustained emission of the vowel “A” by an ASD patient before BoNT-A treatment.

In the present study we aimed to investigate possible differences in voice parameters between ASD patients and HS.

We investigate 20 right-handed native Italian speakers ASD patients and 20 age and sex-matched HS. Symptoms were scored using the Voice Handicap Index scale and a dysphonia clinical scale. Phoniatric evaluation included voice cepstral analysis. The crucial variable in the voice cepstral analysis is the normalized cepstral peak prominence (CPP). We collected voice samples using a high-definition audio recorder (H4n Zoom Corporation, Japan) and a Shure WH20 Dynamic Headset Microphone. Voice samples were digitized at 44.1 kHz, 24
bit, and analysed using the Matlab software. CPP together with other cepstral and spectral features, such as CPPS (smoothed CPP), Hi/Low frequencies rate, harmonics-to-noise ratio, shimmer and jitter were extracted.

Feature selection has been performed by the Waikato Environment for Knowledge Analysis (Weka) software, using Correlation based Feature Selection (CFS) algorithm. The CFS is a filter algorithm that ranks the feature subsets according to a correlation based heuristic evaluation function. The evaluation function prefers subsets that contain features that are highly correlated with the class and uncorrelated with each other. Then, we used a Kononenko’s discretization process of the instances values, that helps the classifier to perform its operations.

Finally, we performed a classification by Weka software, using three different classifiers: Support Vector Machine (SVM), Naïve Bayes, and Multilayer Perceptron Neural Network (ANN).

The proposed method discriminates ASD patients from HS, representing a new helpful tool to better characterize voice abnormalities in ASD [4]. These results suggest the idea that voice features extraction and classification are important instruments to support clinicians in the correct diagnosis of ASD, among different voice disorders.

The results are shown in Table 1, in terms of accuracy, sensitivity, specificity, positive predictive value, and negative predictive value.

<table>
<thead>
<tr>
<th><strong>Classification method</strong></th>
<th><strong>Feature selection</strong></th>
<th><strong>Discretization</strong></th>
<th><strong>Accuracy (%)</strong></th>
<th><strong>Specificity (%)</strong></th>
<th><strong>Sensitivity (%)</strong></th>
<th><strong>NPV (%)</strong></th>
<th><strong>PPV (%)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>No</td>
<td>No</td>
<td>82,22</td>
<td>87,50</td>
<td>76,19</td>
<td>80,77</td>
<td>84,21</td>
</tr>
<tr>
<td>SVM</td>
<td>Yes</td>
<td>Yes</td>
<td>97,78</td>
<td>100,00</td>
<td>95,24</td>
<td>96,00</td>
<td>100,00</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>No</td>
<td>No</td>
<td>82,22</td>
<td>79,17</td>
<td>85,71</td>
<td>86,36</td>
<td>78,26</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>Yes</td>
<td>Yes</td>
<td>97,78</td>
<td>100,00</td>
<td>95,24</td>
<td>96,00</td>
<td>100,00</td>
</tr>
<tr>
<td>ANN</td>
<td>No</td>
<td>No</td>
<td>80,00</td>
<td>83,33</td>
<td>76,19</td>
<td>80,00</td>
<td>80,00</td>
</tr>
<tr>
<td>ANN</td>
<td>Yes</td>
<td>Yes</td>
<td>93,33</td>
<td>100,00</td>
<td>85,71</td>
<td>88,89</td>
<td>100,00</td>
</tr>
</tbody>
</table>

REFERENCES