Automatic Assessment of Voice Signals
According to the GRBAS Scale

Farideh Jalalinajafabadi
The University of Manchester
Manchester, United Kingdom
Jalalinf@cs.man.ac.uk

1. Introduction
Voice disorders may be caused by voice-strain due to speaking or singing, vocal cord damage, infection, side-effects of inhaled steroids as used to treat asthma or more serious disease including laryngeal cancer. The resulting loss of voice quality can be measured subjectively or objectively. For clinical and research use the Japanese Society of Logopeedics and Phoniatrics and the European Research Group recommended a standard referred to as ‘GRBAS’ which is an acronym for a five dimensional scale of measurements of voice properties[1]. The properties are ‘grade’, ‘roughness’, ‘breathiness’, ‘asthenia’ and ‘strain’. The GRBAS scale has the advantage of being widely understood and recommended by many professional bodies, but its subjectivity and reliance on highly trained personnel are significant limitations. The aim of this research is to design and evaluate objective measurement of voice quality conforming to the GRBAS standard.

2. Voice Collection
About 106 voice recordings from patients and volunteers were collected. We have developed a GRBAS presentation and scoring package (GPSP) for the recording of GRBAS scores by speech and language therapists (SLTs) (Figure 2). The GPSP was used to collect and analyse the accuracy and consistency of scoring by three different SLTs.

2.1 GRBAS Scores Analysis
A specific statistical measurement, Cohen’s Kappa coefficient, was used to measure the level of agreement between each pair of SLTs [2]. Figure 3, illustrates moderate agreement between Scorers 1 and 2 for ‘grade’, ‘roughness’, ‘breathiness’ but only fair agreement for ‘asthenia;’ ‘strain’. Scorers 1 and 3 had moderate agreement for all the components of GRBAS, and scorers 2 and 3 had moderate agreement for grade and breathiness and fair agreement for ‘roughness’, ‘asthenia’ and ‘strain’.

3. Voice analysis (DSP) & Conversion to GRBAS
In order to predict the GRBAS scores, the voice features must be extracted. Some parameters produced by the commercial Kay-Pentax software[3], based on ‘MVDP’[4] are intended to be strongly indicative of ‘grade’, ‘roughness’ and ‘breathiness’. A 3-stage classifier was developed as a basic structure to classify the GRBAS scores. This classifier was designed to distinguish, firstly, samples scoring 0 from those scoring 1, 2 or 3, secondly, samples scoring 1 from those scoring 2 or 3, and finally, samples scoring 2 from those scoring 3. Figure 4, shows the average of precision, recall and F-measure.

4. Conclusions
This research shows there is not enough agreement between SLTs in scoring, therefore; GRBAS scores were predicted objectively. To do objective analysis voice features were extracted by Kaypentax and they were used for training the classifier in predicting GRBAS scores. To improve the precision appropriate voice features should be investigated.

References