THE EFFECT OF VISIBLE SPEECH ON PERCEPTUAL RATING OF PATHOLOGICAL VOICES, AND ON CORRELATION BETWEEN PERCEPTION AND ACOUSTICS.

P. H. Dejonckere, J.W.M.A.F. Martens, H. Versnel, M. Moerman
The Institute of Phoniatrics, Department of Otorhinolaryngology, University Medical Centre Utrecht, NL.
COST action 2103 : Advanced Voice Function Assessment

Abstract : The inter-rater variability in perceptual voice evaluation still limits the widespread clinical use of the best available rating system. Support of visible speech in experimental conditions demonstrates a significant enhancement of the inter-rater agreement. However it does not influence the correlation between perceptual and conventional acoustic parameters. The addition of visible speech to the clinical setting is feasible since nowadays affordable computer programs provide the spectrogram in quasi real time.

Keywords : Dysphonia, Perceptual evaluation, Visible Speech, Acoustic analysis.

I. INTRODUCTION

The GIRBAS scale introduced by Hirano [1] has become a commonly used scale for perceptually rating severity of deviance in voice quality. However judgments of different raters (even experienced) might differ considerably [2;3]. Acoustical analysis of pathological voice has several advantages as being quantitative and non-invasive, and cost and time efficient. As a disadvantage, most acoustical analysis relies on quasi-periodic waveforms and thus cannot be used on noisy and irregular voices. Further, because of the lack of one-to-one relations between acoustical and perceptual voice parameters, the perceptual assessment cannot be replaced by acoustical analysis.

Sound spectrograms are another approach: the spectrogram enables visualization of speech (therefore sometimes being referred to as visible speech) and has been widely used in voice and speech research. It was also clinically applied to evaluate voice [4]. The present study examined whether adding visible speech would enhance the interrater agreement of perceptual ratings of pathological voices. Since spectrograms reveal acoustical properties which are related to parameters as jitter and noise-to-harmonic ratio, it is conceivable that visible speech increases the correlations between acoustical and perceptual parameters. Therefore, also the effect of visible speech on these correlations was examined in this study.

II. MATERIALS AND METHODS

Pathological voices
Seventy pathological voices of all kinds of etiologies were digitally recorded. The recorded voice tracks consisted of a prolonged /a/ with a duration of several seconds and a spoken sentence in Dutch.

Visible speech
The visible speech consisted of two spectrograms (0 – 4000 Hz) of the sustained /a/: One spectrogram is produced with a fine frequency resolution (bandwidth: 59 Hz) showing harmonics and the other with a fine time resolution (bandwidth: 300 Hz) showing glottal pulses.

Perceptual evaluation
Six experienced raters independently evaluated the voice samples (prolonged /a/ and sentence, all on a CD) in two sessions with an interval of 4-10 months between the sessions. During the second evaluation session the accessory visible speech of the sustained /a/ was presented to the
Acoustic evaluation
A variety of acoustic parameters as was calculated using the multidimensional voice program (MDVP, Kay Elemetrics Corp.) on a relatively stationary part of the prolonged /a/.

Agreement
Agreement between perceptual evaluations of two experts can be estimated using the parameter kappa (κ) introduced by Cohen. Cohen’s kappa corrects for agreement by chance. To assess the agreement among the six raters we computed kappa according to Fleiss [5] who extended Cohen’s kappa for more than two raters.

To determine whether the agreements found in the conventional and visible-speech conditions significantly differ, the two kappa values were statistically tested.

Acoustical versus perceptual parameters
Since the perceptual parameters (G, I, R, B, A and S) are ordinal, the correlation between the acoustic and perceptual evaluations is calculated using the Spearman rank correlation coefficient.

III. RESULTS

Interrater agreement
The ratings of the 70 voices were used to calculate κ for six raters. The agreement between ratings was significantly higher with than without visible speech for the perceptual parameters G, R and B (Fig. 1).

Acoustical and perceptual parameters
We correlated the perceptual parameters G, I, R, B, A and S with various acoustical parameters for ratings with and without visible speech. No significant changes in correlation were found. Fig. 2 shows the effect for jitter and shimmer.

We investigated the effect of different selection windows on the correlations between acoustical and perceptual parameters. We compared the entire vowel including onset ramp and offset damp, the standard window, and a fixed-duration (1 s) window 250 ms to 1250 ms after onset. These different selection windows did not produce different results on the correlations between acoustical and perceptual parameters.

IV. DISCUSSION

Our study produced two pronounced results. First, the interrater agreement was clearly larger with than without visible speech (information as provided in spectrograms of the voice track) for rating grade, breathiness and roughness. Second, visible speech had no effect on the correlations of the GIRBAS ratings with acoustical parameters.

The addition of visible speech to the clinical setting is feasible since affordable computer programs can provide it in quasi-real-time. Hence, the enhancement of the interrater agreement is an important finding.

No systematic shifts have been found with the addition of visible speech: on the average, G increased whereas B decreased, and R did not shift. Considering the wide distribution of ratings, the ratings with visible speech seem to distinguish well between various voices.

Our results confirm the notion that perceptual rating cannot be replaced by acoustical parameters at least as produced by MDVP paradigms. Perceptual and acoustic measures can be considered complementary. Hence, an optimal evaluation of voice quality is achieved according to a multidimensional protocol, including acoustic and perceptual measures [6;7].

V. CONCLUSION

Support of visible speech demonstrates a significant enhancement of the inter-rater agreement in perceptual voice evaluation. It does not influence the correlation between perceptual and conventional acoustic parameters.
Figure 1. Kappa for 6 raters for G, I, R, B, A and S without and with visible speech. White bars reflect kappa values without visible speech, black bars reflect kappa values with visible speech. Significant differences between kappa with and without visible speech are *: p < 0.05, **: p < 0.001.

Figure 2. Correlations between perceptual parameters grade, roughness and breathiness with the acoustic parameters jitter and shimmer. White bars reflect correlation coefficients without visible speech, black bars reflect correlation coefficients with visible speech. Differences were not significant (p>0.05).
REFERENCES